Gut colonisation by extended-spectrum beta-lactamase-producing Escherichia coli and its association with the gut microbiome and metabolome in Dutch adults: a matched case-control study

Ducarmon, Q.R.; Zwittink, R.D.; Willems, R.P.J.; Verhoeven, A.; Nooij, S.; Klis, F.R.M. van der; ... ; Kuijper, E.J.

Gut colonisation by extended-spectrum β-lactamase-producing *Escherichia coli* and its association with the gut microbiome and metabolome in Dutch adults: a matched case-control study

Quinten R Ducarmon, Romy D Zwittink, Roel P J Willems, Aswin Verhoeven, Sam Nooij, Fiona R M van der Klis, Eelco Franz, Jolanda Kool, Martin Giera, Christina M J E Vandenbroucke-Grauls, Susana Fuentes, Ed J Kuijper

Summary

Background Gut colonisation by extended-spectrum β-lactamase (ESBL)-producing *Escherichia coli* is a risk factor for developing overt infection. The gut microbiome can provide colonisation resistance against enteropathogens, but it remains unclear whether it confers resistance against ESBL-producing *E coli*. We aimed to identify a potential role of the microbiome in controlling colonisation by this antibiotic-resistant bacterium.

Methods For this matched case-control study, we used faeces from 2751 individuals in a Dutch cross-sectional population study (PIENTER-3) to culture ESBL-producing bacteria. Of these, we selected 49 samples that were positive for an ESBL-producing *E coli* (ESBL-positive) and negative for several variables known to affect microbiome composition. These samples were matched 1:1 to ESBL-negative samples on the basis of individuals' age, sex, having been abroad or not in the past 6 months, and ethnicity. Shotgun metagenomic sequencing was done and taxonomic species composition and functional annotations (ie, microbial metabolism and carbohydrate-active enzymes) were determined. Targeted quantitative metabolic profiling (proton nuclear magnetic resonance spectroscopy) was done to investigate metabolomic profiles and combinations of univariate (*t* test and Wilcoxon test), multivariate (principal coordinates analysis, permutational multivariate analysis of variance, and partial least-squares discriminant analysis) and machine-learning approaches (least absolute shrinkage and selection operator and random forests) were used to analyse all the molecular data.

Findings No differences in diversity parameters or in relative abundance were observed between ESBL-positive and ESBL-negative groups based on bacterial species-level composition. Machine-learning approaches using microbiota composition did not accurately predict ESBL status (area under the receiver operating characteristic curve [AUROC]=0·41) when using either microbiota composition or any of the functional profiles. The metabolome also did not differ between ESBL groups, as assessed by various methods including random forest (AUROC=0·61).

Interpretation By combining multiomics and machine-learning approaches, we conclude that asymptomatic gut carriage of ESBL-producing *E coli* is not associated with an altered microbiome composition or function. This finding might suggest that microbiome-mediated colonisation resistance against ESBL-producing *E coli* is not as relevant as it is against other enteropathogens and antibiotic-resistant bacteria.

Funding None.

Lancet Microbe 2022; 3: e443–51

Published Online April 20, 2022 https://doi.org/10.1016/S2666-5247(22)00037-4

Department of Medical Microbiology (Q R Ducarmon PhD, R D Zwittink PhD, S Nooij MSc, Prof E J Kuijper PhD) **and Centre for Microbiome Analyses and Therapeutics** (Q R Ducarmon, R D Zwittink, S Nooij, Prof E J Kuijper) **and Centre for Proteomics and Metabolomics** (A Verhoeven PhD, M Giera PhD), **Leiden University Medical Centre, Leiden, Netherlands; Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centres, Amsterdam Medical Centre, Amsterdam Infection and Immunity Institute, Amsterdam, Netherlands** (R P J Willems MD, Prof C M J E Vandenbroucke-Grauls PhD)**; Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, Netherlands** (F R M van der Klis PhD, E Franz PhD, J Kool BSc, S Fuentes PhD, Prof E J Kuijper)

Correspondence to: Dr Quinten R Ducarmon, Department of Medical Microbiology, Leiden University Medical Centre, 2333ZA Leiden, Netherlands **q.r.ducarmon@lumc.nl**

Introduction

Escherichia coli is a common gut commensal, but several strains possess virulence factors that enable them to cause gastrointestinal, urinary, and extraintestinal infections.[1,2] Colonisation of the gut by antibiotic-resistant organisms, including extended-spectrum β-lactamase (ESBL)-producing *E coli* and carbapenem-resistant *E coli*, often precedes infections.[3] ESBL-producing *E coli* pose a great threat to public health, with approximately 4900 reported bloodstream infections annually in England caused by this bacterium.[4] The US Centers for Disease Control and Prevention marked ESBL-producing Enterobacteriaceae as a serious threat to public health.[5] In addition, the prevalence of clinical *E coli* isolates that produce ESBL has strongly increased in the past few decades and *E coli* is the leading causative pathogen in Gram-negative bacteraemia.[6,7] The gut microbiome can mediate colonisation resistance against several enteric pathogens, but it remains unclear whether this is also the case for antibiotic-resistant bacteria such as ESBL-producing *E coli*, especially because many individuals harbour commensal *E coli*. Colonisation resistance can be conferred by the gut microbiome through nutrient competition, production of antimicrobial compounds,

Research in context

Evidence before this study

We searched PubMed for research articles published in English from database inception up to Dec 10, 2021, using the terms ("Gut microbiome" OR "Gut microbiota") AND ("Colonisation resistance" OR "Colonisation resistance") AND ("ESBL" OR "MDRO") AND ("Humans" OR "Volunteers" OR "Residents" OR "Participants" OR "Patients") AND ("Enterobacterales" OR "Enterobacteriaceae" OR "*Escherichia coli*" OR "*E. coli*") NOT ("Review"[publication type]). This search identified 33 articles, of which ten articles investigated the relation between the gut microbiome and the provision of colonisation resistance against ESBL-producing Enterobacterales or other multidrug-resistant organisms in humans. These ten studies were done in various populations (eg, residents of nursing homes and patients receiving haematological treatment) and were generally restricted to profiling the microbiota using 16S ribosomal ribonucleic acid gene amplicon sequencing, with no information on metabolomics. Also, none of these studies incorporated case-control matching into their study. Therefore, the effects of confounders such as age, underlying diseases, or medication use on microbiome composition could not be excluded. This fact is likely to be an important reason why previous studies have yielded conflicting evidence regarding the potential role of the gut microbiome in providing colonisation resistance against extended-spectrum β-lactamase (ESBL)-producing bacteria (and specifically, *Escherichia coli*). Therefore, it currently remains unknown whether or not the gut microbiome can provide colonisation resistance against ESBL-producing *E coli*.

Added value of this study

To our knowledge, this is the first study to investigate microbiome-mediated colonisation resistance against ESBL-producing *E coli* using a combination of metagenomics, metabolomics, and machine-learning approaches.
In addition, this study applied stringent inclusion criteria, and individuals colonised by an ESBL-producing *E coli* were carefully matched to non-colonised individuals on the basis of several epidemiological and clinical variables.
Also, we focused our study on ESBL-producing *E coli*, whereas previous studies generally analysed potentially protective bacterial taxa against a broad range of antibiotic-resistant bacteria. The latter approach is less sensitive, owing to large genomic and functional differences between bacterial species.
We found that there were no differences in the gut microbiome (at both the taxonomic and functional level) or metabolome of individuals who are asymptomatically colonised by an ESBL-producing *E coli* when compared with matched non-colonised individuals.

Implications of all the available evidence

These findings suggest that microbiome-mediated colonisation resistance might not be as relevant against ESBL-producing *E coli* as it is for other Gram-positive enteric pathogens (eg, *Clostridioides difficile* and vancomycin-resistant *Enterococcus*). Therefore, microbiome-based interventions might not be the way forward to prevent or eradicate intestinal colonisation of ESBL-producing *E coli*.

support of gut barrier integrity, bacteriophage deployment, and through interaction with the immune system.[8] However, studies in humans have reported conflicting evidence regarding which bacterial genera or species within the gut microbiome could be of relevance in providing colonisation resistance against ESBL-producing *E coli* or ESBL-producing Enterobacterales. These conflicting results can, at least partly, be traced back to several confounding factors (eg, medication and comorbidities) in those studies.[9–12] It was shown that unevenly matched case-control studies with regard to lifestyle and physiological characteristics can produce spurious microbial associations with human-phenotypes-like disease, or in this case, colonisation by ESBL-producing *E coli*.[13] Therefore, if increasingly robust signatures related to colonisation resistance could be identified, this could pave the way for novel, urgently needed microbiome-based therapeutics against this antibiotic-resistant bacterium.

Here, we aimed to compare the gut microbiome and metabolome between individuals asymptomatically colonised with an ESBL-producing *E coli* (ESBL-positive) and individuals who were not (ESBL-negative), as determined by culture-based and molecular approaches, to identify a possible role of the microbiome in the control of colonisation by this antibiotic-resistant bacterium. To avoid confounding factors from affecting study results, we used data from a large Dutch cross-sectional population study (PIENTER-3).[14] From this cohort, faecal samples of 2751 individuals (one faecal sample per individual) were used to culture ESBL-producing bacteria. With this high number of samples available, we could apply stringent sample selection with regard to known confounders in microbiome studies, such as antibiotic use, proton-pump inhibitor use, and various diets. Subsequently, we did case-control matching using various epidemiological and health-related variables. We did extensive functional and taxonomic profiling of the gut microbiome through metagenomics and metabolomics, to investigate whether there are differences in the gut microbiome between matched ESBL-positive and ESBL-negative samples.

Methods

Sample collection and selection

For our matched case-control study, samples were selected from a larger Dutch population-wide study (PIENTER-3).[14] This cross-sectional population study was done between Feb 1, 2016, and Oct 16, 2017, and was primarily designed to obtain insight into age-specific seroprevalence of vaccine-preventable infectious diseases.

The study proposal was approved by the Medical Ethics Committee Noord-Holland (METC number M015–022) and written informed consent was obtained from all participants or their legal guardian.[14]

For our study we ultimately included 98 samples, of which 95 were stored in the freezer within 15 min after defecation, one person did not provide information on this, and two individuals took longer than 1 h to store their sample in the freezer. Samples were kept on average for 2·97 days (SD 2·82; six individuals did not indicate this information) in people's freezers before being delivered on cold packs to the mobile study team.[14] Faecal samples were kept on dry ice during transport to the Dutch National Institute for Public Health and the Environment and stored at –80°C from the next day.

We took the following steps. First, 2751 faecal samples were cultured for ESBL-producing bacteria, of which 198 samples were positive. We then selected samples that were positive for ESBL-producing *E coli*, resulting in 176 potential samples. Next, for all samples we applied stringent exclusion criteria based on variables known to affect the gut microbiome. Individuals were excluded if they met the following criteria: current proton-pump inhibitor use, antibiotic use in the past 3 months, diarrhoeal symptoms (defined as at least three thin stools within 24 h) in the past month, vomiting in the past month, blood in stool during the past month, abdominal pain or nausea during the past month, use of any prebiotics or probiotics, consumption of a special diet (ie, a vegetarian diet; a diet free of cow milk, hen-egg protein, gluten, nuts, peanuts, or colouring agents; a diet with restricted lactose, protein, fat, cholesterol, calories, or salt; a diabetes-related diet; or a diet enriched with dietary fibre, energy, or protein, or that is easily digestible; or any other special diet) and whether stool was stored in the freezer after defecation (samples were excluded if not stored in the freezer). This selection resulted in 51 ESBL-positive samples for inclusion, which were subsequently matched to 51 ESBL-negative samples using the R MatchIt software package (version 3.0.2), using the nearest method in the MatchIt function. Patients were matched for age, sex, having been abroad or not during the past 6 months, and ethnicity. Three samples (one ESBL-negative sample and two ESBL-positive samples) were excluded because insufficient DNA was available for sequencing. One additional sample (ESBL-negative) was excluded because after sequencing, we discovered that this individual had provided ambiguous answers regarding dietary habits. The final dataset for analysis contained 49 individuals in each group.

Detection of ESBL-producing Enterobacterales

Details of the microbiological methods have been described elsewhere[15] and are summarised in the appendix (p 2).

DNA extraction for metagenomic shotgun sequencing

All information regarding DNA extraction from the whole stool can be found in the appendix (p 2).

Metagenomic shotgun sequencing

Shotgun metagenomic sequencing was done by GenomeScan (Leiden, the Netherlands) using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) and the Illumina NextSeq 500 platform (San Diego, CA, USA; paired-end, 150 base pairs). Two positive sequencing controls (ZymoBiomics Microbial Community DNA Standards [Zymo Research, Irvine, CA, USA]) and two negative sequencing controls (sterile water) were included. The mean number of raw reads was 4747908 (range 2565232–62035096) and the median was 4142237 paired-end reads. All details about processing metagenomic sequencing data and statistical analyses can be found in the appendix (pp 2–3).

Metabolomics

We did metabolomic analysis using proton nuclear magnetic resonance, as metabolomics can be seen as a functional readout of the microbiome and is complementary to functional prediction of metagenomes. The method for proton nuclear magnetic resonance analysis of faecal samples used the protocol developed by Kim and colleagues,[16] with a few minor adaptations (appendix p 3). All details about measurements of metabolites and subsequent processing can be found in the appendix (p 3).

Role of the funding source

There was no funding source for this study.

Results

Demographic and participant characteristics were highly similar between the ESBL-positive and ESBL-negative groups (n=49 individuals in each group) and antibiotic use in the preceding 3–12 months was also evenly matched (table). Regarding the ESBL-producing *E coli* isolates that colonised our 49 ESBL-positive participants, 44 people carried a CTX-M type (as determined by PCR). The majority of these were CTX-M-1 (n=25) and CTX-M-9 (n=18) and one could not definitively be typed (CTX-M-1 or CTX-M-8). Isolates of four individuals were negative for *CTX-M* genes and for one participant it could not be determined. Additional information on antimicrobial susceptibility of the strains can be found in the appendix (p 5).

We first investigated potential differences in microbiota composition and diversity between ESBL-positive and ESBL-negative samples. A total of 1178 species (ie, metagenomic operational taxonomic units [mOTUs]) were detected in our cohort. Overall bacterial composition at the family and genus level is shown in the appendix (p 6). The most abundant species in this cohort were *Bifidobacterium adolescentis* (mean relative abundance 4·6% [SD 6·9%]), *Ruminococcus bromii* (3·4% [4·8%]), undefined

See Online for appendix

	ESBL-negative (n=49)	ESBL-positive (n=49)	p value*
Age, years			
Mean (SD)	44·1 (15·2)	46·6 (15·3)	0·43
Median (IQR)	45·0 (28–55)	46·0 (36–59)	..
Sex			
Male	26 (53%)	23 (47%)	0·69
Female	23 (47%)	26 (53%)	..
Travelled abroad in the past 6 months			
Yes	39 (80%)	37 (76%)	0·81
No	10 (20%)	12 (24%)	..
Ethnicity			
Dutch	38 (78%)	36 (73%)	0·79
First generation other western	1 (2%)	0 (0%)	..
Second generation other western	2 (4%)	3 (6%)	..
First generation Suriname, Aruba, or Dutch Antilles	3 (6%)	3 (6%)	..
Second generation Suriname, Aruba, or Dutch Antilles	1 (2%)	0 (0%)	..
First generation other non-western	4 (8%)	7 (14%)	..
Antibiotics used in the past 3-12 months			
Yes	6 (12%)	7 (14%)	0·77
No	43 (88%)	41 (84%)	..
Do not know	0 (0%)	1 (2%)	..

Data are n (%), mean (SD), or median (IQR). We used the Dutch Central Bureau of Statistics definitions of western and non-western; western denotes Europe (except Turkey), North-America, Oceania, Indonesia, and Japan, whereas non-western denotes Asia (except Indonesia and Japan), Latin America, Africa, and Turkey. ESBL=extended-spectrum β-lactamase. *p values were obtained using an independent *t* test for continuous data variables and Fisher's exact test for categorical variables.

Table: **Participant characteristics**

members of the family Ruminococcaceae (2·9% [3·2%]), *Eubacterium rectale* (2·7% [2·8%]), and *Prevotella copri* (2·5% [5·7%]). We did not observe differences in alpha diversity (ie, observed mOTUs and Shannon index; figure 1A, B), or beta diversity (principal coordinates analysis and *t*-distributed stochastic neighbour embedding; figure 1C, D). We computed Mash distances to complement our analyses based on marker genes that were done using the mOTUs profiler, but ordination looked very similar to figure 1D, without a clear group separation (appendix p 6).

Next, we investigated whether there were differences in relative abundance between the study groups at the species level (mOTUs). Before differential abundance testing, mOTUs were filtered on the basis of a prevalence of at least 10%, resulting in 436 mOTUs (representing 37% of the total observed mOTUs). No significantly different abundances of mOTUs were detected (all corrected p values >0·81). To see whether microbiota composition was predictive of ESBL-producing *E coli* carriage, we applied a machine-learning classifier (least absolute shrinkage and selection operator logistic regression) to the filtered mOTUs relative abundance matrix, which provided an area under the receiver operating characteristic curve (AUROC) value of 0·41. This value is slightly worse than random classification, although the 95% CI included an AUROC of 0·5 (which indicates random classification) throughout the figure (figure 1E), which indicates that the relative abundance of mOTUs did not allow for reliable prediction of ESBL status. In conclusion, no differences between ESBL-positive and ESBL-negative samples were found in bacterial species composition or diversity parameters.

Next, we compared the resistomes of ESBL-positive and ESBL-negative samples. Of all cleaned reads (ie, the reads that remain after quality filtering and the removal of human reads), a mean of 0·035% (SD 0·024%) reads per sample mapped against the MegaRes 2.0 database.[17] There was no difference between ESBL groups in the average number of reads aligned to MegaRes 2.0 (independent *t* test p=0·84). A total of 98 unique antimicrobial resistance genes (ARGs) were detected, with 17 different mechanisms of antimicrobial resistance (eg, β-lactam), and the number of detected ARGs was not different between ESBL groups (independent *t* test p=0·46; figure 2A). Overall, ARG profiles in the study groups assessed by plotting beta diversity did not show a clear separation between ESBL groups (figure 2B), which was confirmed by permutational multivariate analysis of variance (p=0·21). The most abundant ARGs and mechanisms of antimicrobial resistance are shown in figure 2C and D. No differences in relative abundance of ARGs were found between the groups using differential abundance analysis (all corrected p values >0·44). Tetracycline resistance was most abundant in the resistomes (mean 47·7% [SD 24·7%]; figure 2C), followed by mupirocin resistance (33·7% [28·6%]). Tetracycline resistance was conferred by several *tet* genes, whereas mupirocin resistance was conferred through the *ileS* gene. Because it is known from the literature that *Bifidobacterium* spp can be intrinsically resistant to mupirocin through the *ileS* gene,[18] we analysed the correlation between the relative abundance of *Bifidobacterium* (at genus level) and the *ileS* gene, which was indeed high (R=0·78; $p<2·2\times10^{16}$; appendix p 7). We then investigated the functional profiles of our participants.

To compare the functionality of the gut microbiome between the study groups, cleaned reads were mapped against the annotated integrated gene catalog. A mean of 95·8% (SD 1·7%) of reads aligned with the integrated gene catalog, and the aligned number of reads was not different between ESBL groups (independent *t* test p=0·23). From the aligned reads, 49·2% (SD 2·2%) aligned against a gene annotated by a functional group (KEGG orthology [KO] group) and this was not different between ESBL groups (independent *t* test p=0·13). There was no difference in overall functional profiles between

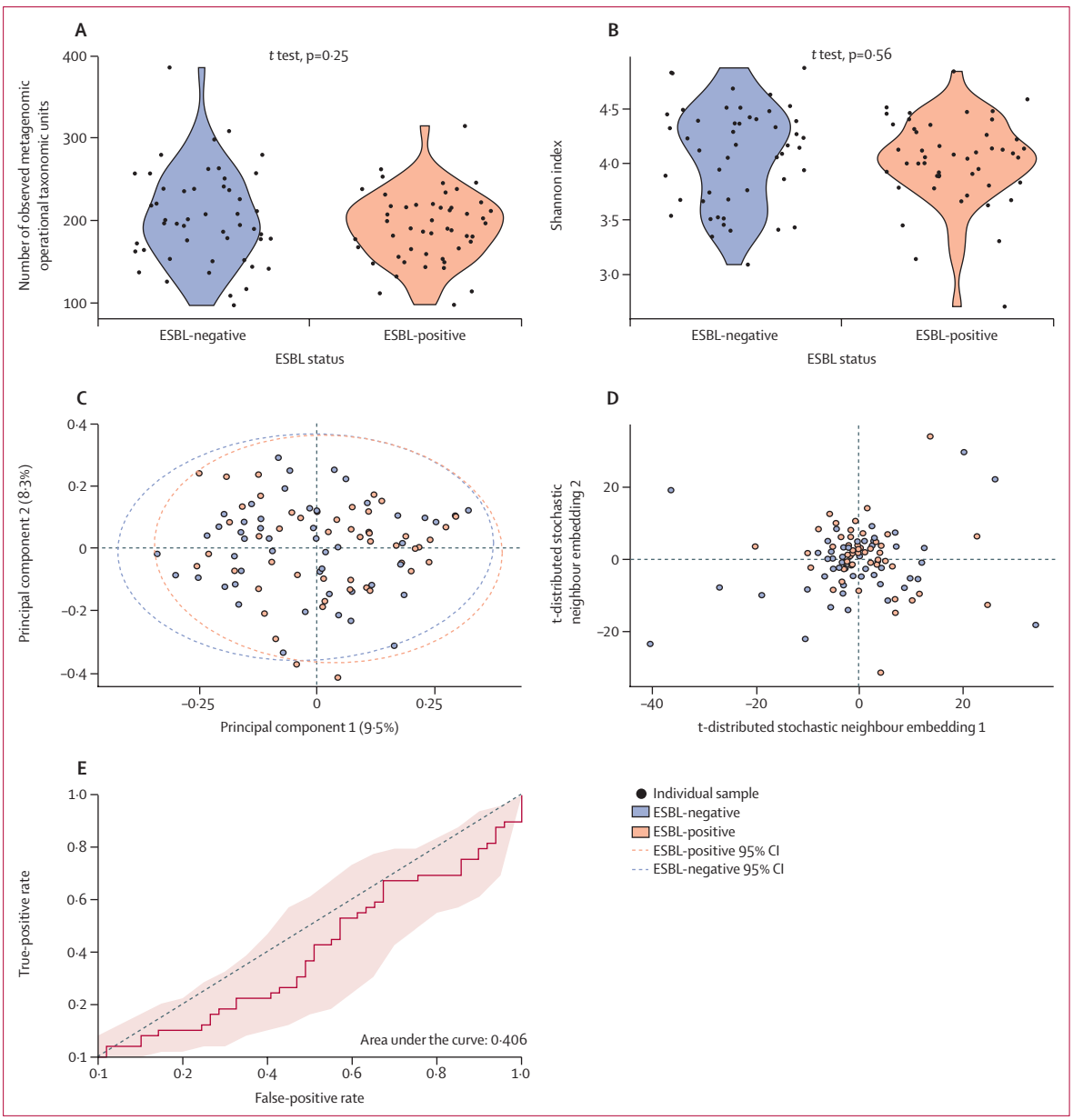

***Figure 1*: Taxonomic analyses at bacterial species level**

(A) Comparison of the observed metagenomic operational taxonomic units. (B) Comparisons of the Shannon index. (C) Unsupervised clustering using principal coordinates analysis based on Bray–Curtis dissimilarity. (D) t-distributed stochastic neighbour embedding based on Bray–Curtis dissimilarity. (E) The receiver operating characteristic curve for the least absolute shrinkage and selection operator on the metagenomic operational taxonomic units relative abundance, with the mean area under the curve value (red line) and its respective 95% CI (red area). ESBL=extended-spectrum β-lactamase.

the groups (permutational multivariate analysis of variance p=0·19). 8450 KO groups were detected and after filtering on 10% prevalence, 5987 KO groups remained for differential abundance testing. No KO groups were significantly differentially abundant between ESBL groups (all corrected p values >0·25). To identify functional groups predictive of ESBL status, least absolute shrinkage and selection operator logistic regression was applied to the relative abundance matrix of KO groups. No accurate prediction model could be constructed (AUROC of 0·60), indicating that the functional groups did not contain information allowing for prediction of ESBL status. Taken together, these considerations led to the conclusion that there was no difference in encoded microbiome functionality between the groups.

Next, the carbohydrate-active enzymes (CAZymes) were profiled to investigate potential differences in carbohydrate metabolism. From the aligned reads, a mean of 2·1% (SD 0·2%) aligned against a gene annotated to a CAZyme family, and this was not different

between ESBL groups (independent *t* test p=0·48). A total of 109 CAZyme families were detected with a mean of 77·7 (SD 5·7) per individual, with no differences between ESBL groups (independent *t* test p=0·34; figure 3A). The three most abundant CAZymes in our study were glycoside hydrolase (GH)13 (19·4% [SD 3·3%]), GH3 (11·4% [1·6%]), and GH31 (6·2% [0·9%]; figure 3C), corresponding to the breakdown of starch and glycogen (GH13) and the breakdown of plant cell wall glycans (GH3 and GH31).[19] Variation in CAZyme relative abundance profiles could not be explained by ESBL group (permutational multivariate analysis of variance p=0·57; figure 3B). Compositional plots based on the 20 most abundant CAZymes were highly similar between the ESBL groups (figure 3C), and no differences in relative abundance of individual CAZyme families were observed (all corrected p values >0·71). To identify potential drivers of ESBL-producing *E coli* colonisation, we used least absolute shrinkage and selection operator logistic regression on relative abundances of CAZymes, which did not result in an accurate prediction model (AUROC 0·56). This indicated there was only very low to no predictive power in relative abundances of CAZymes with regard to ESBL status.

For metabolomic analysis we quantified metabolite concentrations in all individuals, except for one ESBL-positive sample that was excluded as a good quality proton nuclear magnetic resonance spectrum could not be recorded because of problems with the shimming process used to adjust the instrument. First, to investigate whether any differences in metabolite concentrations existed between ESBL groups, we did univariate testing using independent *t* tests. These results strongly depended on the method used for multi-error correction (11 metabolites were significantly different at p=0·048 with the Benjamini-Hochberg procedure, but none were significantly different with the Holm procedure; appendix p 8–9).

Unsupervised dimensionality reduction using principal component analysis was done to investigate whether, based on metabolome profiles, any separation could be observed between ESBL groups (figure 4A). More than 46% of the metabolome variation could be explained on the first principal component, with some separation of the study groups. However, supervised analysis using a partial least-squares discriminant analysis indicates that no predictive value could be obtained for class separation based on two partial least-squares components (Q2Y –0·06). Lastly, we ran a random forest prediction model to investigate whether ESBL status could be predicted on the basis of metabolite profiles, but it could not (AUROC 0·61; figure 4B). Altogether, minor differences in metabolite concentrations could be detected using *t* tests, but these were dependent on the method applied for correction for

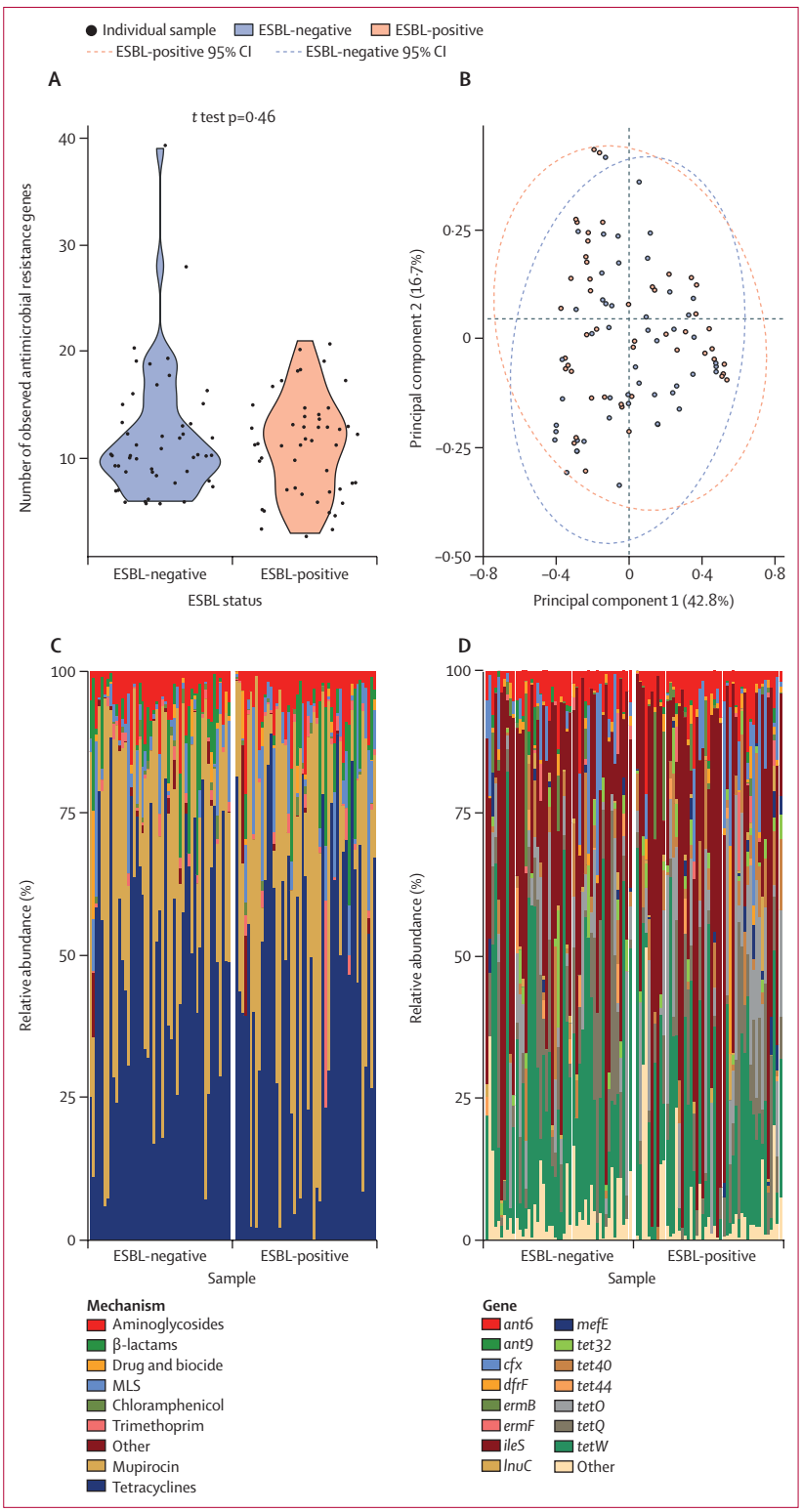

Figure 2: **Resistome analyse**
Comparison of the number of detected antimicrobial resistance genes (A). Principal coordinates analysis of antimicrobial resistance genes based on Bray–Curtis dissimilarity (B). Compositional plots of the most abundant resistance mechanisms (C) and genes (D). ESBL=extended-spectrum β-lactamase. MLS=macrolides, lincosamides, and streptrogramin A and B antibiotics.

multiple testing. Principal component analysis between the ESBL groups showed, at best, a small overall signal, but no predictive value could be confirmed by either partial least-squares discriminant analysis or random forest modelling. In conclusion, metabolomics profiling did not reveal any differences between the groups.

Discussion

We present a unique study that investigated differences in the gut microbiome and metabolome between individuals asymptomatically colonised by an ESBL-producing *E coli* and matched non-colonised individuals. By contrast with previous studies on this topic, we applied stringent inclusion criteria and matched ESBL-positive individuals with ESBL-negative individuals on important epidemiological variables, which minimised the chance for observing effects that could be attributed to confounding variables. The combination of metagenomics and metabolomics allowed for a deep molecular resolution of the gut microbiome at the taxonomic and functional level. We showed that there is no difference in the gut microbiome of individuals asymptomatically colonised with an ESBL-producing *E coli* as compared with non-colonised individuals.

Confounding factors might be the reason for the previously reported differences in microbial signatures associated with protection from asymptomatic colonisation by ESBL-producing bacteria and multidrug-resistant organisms across different studies. It must be noted that these studies have mostly investigated patient populations that are susceptible, such as residents in nursing homes and patients in hospital. In such populations, it is very complex to disentangle observed differences between colonised and non-colonised individuals from the differences caused by confounding variables (such as comorbidities and medication) in compared individuals.[10,12,20–22] In our study we excluded individuals on the basis of many microbiome-influencing clinical factors, and did matching on several clinical variables, as recommended for cross-sectional microbiome studies.[13] In this way, we could study the effect of colonisation of ESBL-producing *E coli* in isolation and convincingly show that no differences exist in the gut microbiome between colonised and non-colonised individuals.

Previous research has generally not focused on species-specific colonisation resistance, but rather on a broad category of multidrug-resistant organisms (such as ESBL-producing Enterobacterales).[10,12,20–22] Given the large genomic diversity within species,[23] let alone within the order of Enterobacterales, it is highly unlikely that a common mechanism exists that could prevent colonisation of, for example, both ESBL-producing *Klebsiella pneumoniae* and ESBL-producing *E coli*. Therefore, in the current study, we focused on a single species (*E coli*), rather than on a broad group of ESBL-producing Enterobacterales.

The microbiome composition of individuals in our study population reflected that of other population cohorts in general. For example, *B adolescentis* has been previously described in another Dutch cohort[24] as the most abundant bacterial species, with an average relative abundance of 9·51% (SD 10·8%). In addition, *P copri*, *R bromii*, and *E rectale* were also highly abundant and prevalent, in line with the findings in the current study.[24]

The resistome profiles identified in our study also corresponded well with what is generally described in literature, with tetracycline resistance being the most abundant resistance mechanism in the human gut.[25] The observed high relative abundance of *ileS* (associated with mupirocin resistance) in the metagenomes of our study participants could be explained by the high relative abundances of *Bifidobacterium* spp in this cohort.

We show that despite interindividual variation in taxonomic profiles, the functionality of the microbiome,

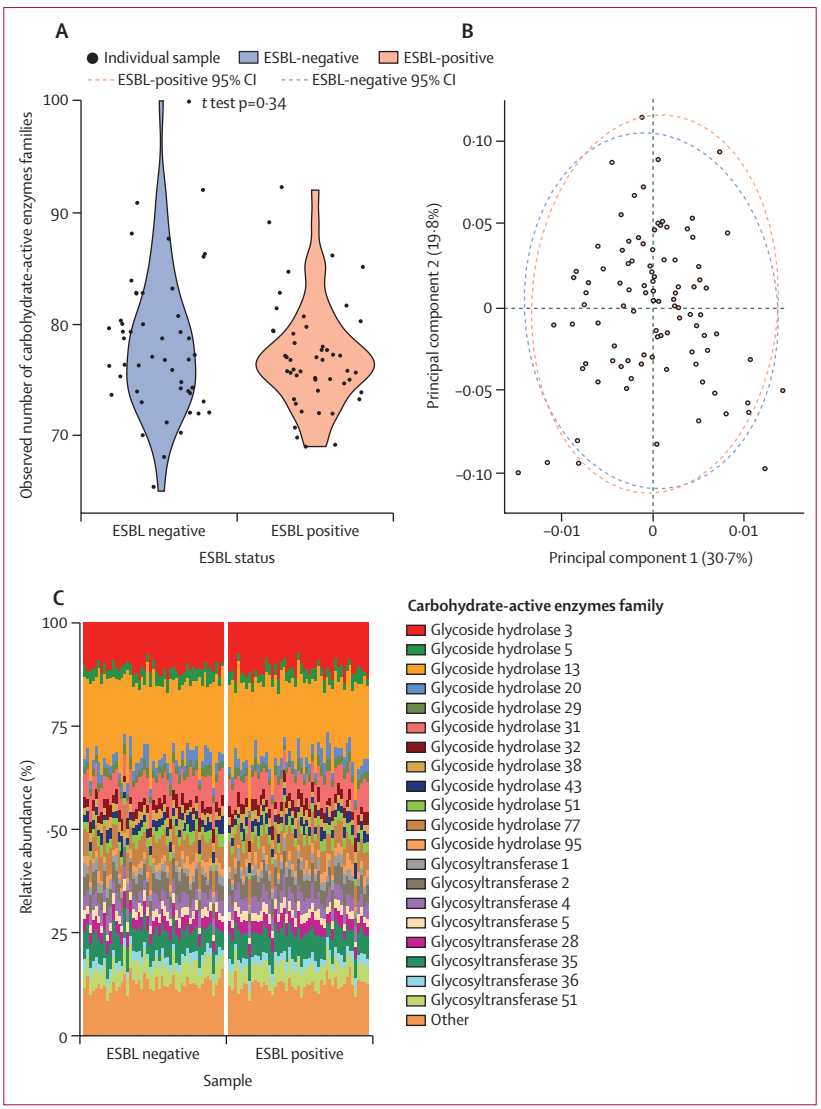

Figure 3: **Analyses based on carbohydrate-active enzymes repertoire**
(A) Comparison of the number of carbohydrate-active enzyme families. (B) Principal coordinates analysis of carbohydrate-active enzymes based on Bray–Curtis dissimilarity. (C) Compositional plot of the consistency of carbohydrate-active enzymes families among participants. ESBL=extended-spectrum β-lactamase.

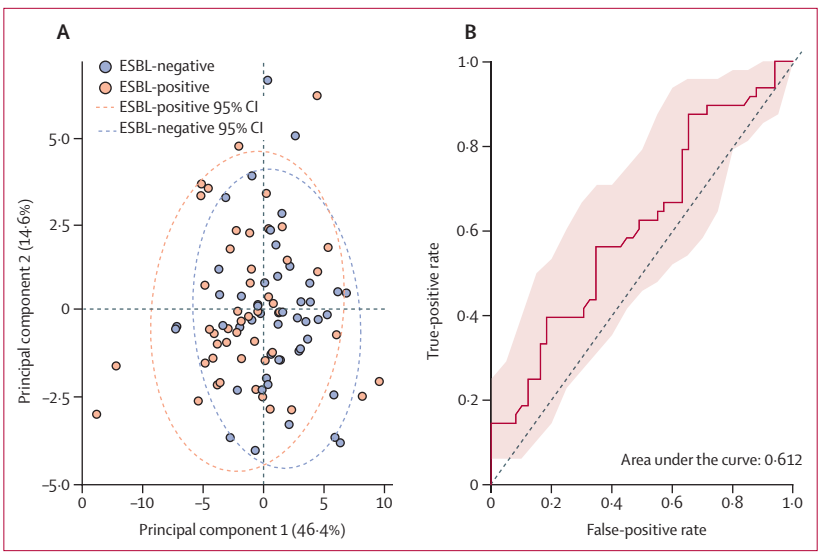

Figure 4: **Metabolomic analyses**
(A) Principal component analysis. (B) The receiver operating characteristic curve for random forest analysis based on metabolite concentrations, with the mean area under the curve value (red line) and its respective 95% CI (pink area). ESBL=extended-spectrum β-lactamase.

as assessed by, for example, the relative abundance of CAZyme families, is highly consistent between individuals. These finding are in line with previous findings showing functional similarity at the metabolic level despite taxonomic diversity.[26]

This study is, to our knowledge, the first study to profile the gut metabolome in relation to colonisation of ESBL-producing *E coli*. We did not observe a relation between the metabolome, or any specific metabolite, and ESBL status. For enteric pathogens like *Salmonella enterica* serovar Typhimurium and *Clostridioides difficile*, specific metabolites have been shown to be strongly related to colonisation resistance in rodent models.[27,28] Of note, these two studies[27,28] were infection models, whereas colonisation models would have better represented our study.

A limitation of our study is the absence of longitudinal microbiome data for the participants, which means that we cannot make any statements about the duration of colonisation of ESBL-producing *E coli* and associations with the gut microbiome over time. This issue is particularly relevant considering the large variation in colonisation duration between individuals.[29,30] It could be speculated that individuals who are long-term colonised have a different gut microbiome compared with individuals who are only colonised for a short period of time, although there is no clear evidence for this in the literature, to our knowledge. Furthermore, longitudinal observations would allow us to identify potential changes occurring at the compositional and functional levels when asymptomatic carriage turns into active infection or when people become decolonised. A second limitation of our study is that we could not estimate the abundance of ESBL-producing *E coli* in comparison to the total *E coli* population. Such an estimate could be investigated by constructing metagenome-assembled genomes of different *E coli* strains within the same sample, but this was not feasible because of the relatively low sequencing depth in combination with the very low abundance of *E coli* in our samples. Third, ideally a study would have microbiome data for an individual both from shortly before an ESBL-producing *E coli* had colonised and at the time of colonisation, so that microbiome changes within an individual could be investigated. Last, we did not have whole-genome sequencing data of the ESBL-producing *E coli* isolates, which prevented us from placing these data into a broader epidemiological context or from making any claims about their potential pathogenicity.

However, this study is unique in that ESBL-positive and ESBL-negative individuals were selected from a large Dutch cohort (n=2751), and therefore we could apply stringent inclusion criteria and match the two groups on several demographic and clinical variables. To our knowledge, this is one of very few studies in the microbiome field that applied such a stringent study setup, which ensured that the potential effect of confounding factors was minimised. Also, to our knowledge, this study is the first to investigate differences in the gut microbiome and metabolome between individuals colonised by an ESBL-producing *E coli* and non-colonised individuals, using a combined approach of metagenomics and metabolomics, and therefore provides insights into both the composition and the function of the gut microbiome.

In conclusion, our study does not show differences in the gut microbiome or metabolome of individuals who are, or who are not, colonised by an ESBL-producing *E coli*. We hypothesise that microbiome-mediated colonisation resistance might therefore not be as relevant against ESBL-producing *E coli* as it is for other enteric pathogens (like *C difficile* and vancomycin-resistant *Enterococcus*), although longitudinal studies or controlled human colonisation models are necessary to confirm this hypothesis.

Contributors

QRD, RDZ, and EJK conceived and designed the study. RDZ, SF, and EJK supervised the study. RPJW and CMJEV-G did the study on ESBL-producing bacteria in the open population in the Netherlands and analysed all patient-related data. AV and MG did the metabolomics analysis and AV aided in statistical analysis of metabolomics. FRMvdK and EF coordinated sample collection. JK did the DNA extraction and related laboratory procedures. QRD did sample selection and analysed epidemiological data, processed and analysed metagenomic data, did statistical analysis of metabolomics data, created the figures, and wrote the manuscript. SN did the Mash analysis. All authors discussed the results and implications. QRD and SF accessed and verified the data. All authors had full access to all the data in the study, contributed to and approved the manuscript, and had final responsibility for the decision to submit for publication.

Declaration of interests

EJK is supported by an unrestricted grant from Vedanta Biosciences, has done research for Cubist, Novartis, and Qiagen, and has participated in

advisory forums for Astellas, Optimer, Actelion, Pfizer, Sanofi Pasteur, and Seres Therapeutics, all outside of this work. All other authors declare no competing interests. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data sharing

All raw metagenomic data can be found under project PRJEB44119 on the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/view/PRJEB44119?show=reads). All R code and necessary data files to reproduce the analyses and figures have been uploaded to GitHub (https://github.com/qducarmon/ESBL-Ecoli_Colonization_Microbiome_Metabolome).

Acknowledgments

We gratefully acknowledge all the individuals who participated in the PIENTER-3 study.

References

1 Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin Microbiol Rev* 2013; **26:** 822–80.

2 Pitout JDD, Laupland KB. Extended-spectrum beta-lactamase-producing Enterobacteriaceae: an emerging public-health concern. *Lancet Infect Dis* 2008; **8:** 159–66.

3 Cassone M, Mody L. Colonisation with multi-drug resistant organisms in nursing homes: scope, importance, and management. *Curr Geriatr Rep* 2015; **4:** 87–95.

4 Day MJ, Hopkins KL, Wareham DW, et al. Extended-spectrum β-lactamase-producing *Escherichia coli* in human-derived and foodchain-derived samples from England, Wales, and Scotland: an epidemiological surveillance and typing study. *Lancet Infect Dis* 2019; **19:** 1325–35.

5 US Department of Health and Human Services, Centres for Disease Control and Prevention. Antibiotic resistance threats in the United States 2019. December, 2019. https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf (accessed Dec 10, 2021).

6 Denisuik AJ, Karlowsky JA, Adam HJ, et al. Dramatic rise in the proportion of ESBL-producing *Escherichia coli* and *Klebsiella pneumoniae* among clinical isolates identified in Canadian hospital laboratories from 2007 to 2016. *J Antimicrob Chemother* 2019; **74** (suppl 4)**:** iv64–71.

7 Vihta KD, Stoesser N, Llewelyn MJ, et al. Trends over time in *Escherichia coli* bloodstream infections, urinary tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998–2016: a study of electronic health records. *Lancet Infect Dis* 2018; **18:** 1138–49.

8 Ducarmon QR, Zwittink RD, Hornung BVH, van Schaik W, Young VB, Kuijper EJ. Gut microbiota and colonisation resistance against bacterial enteric infection. *Microbiol Mol Biol Rev* 2019; **83:** e00007-19.

9 Huang YS, Lai LC, Chen YA, et al. Colonisation with multidrug-resistant organisms among healthy adults in the community setting: prevalence, risk factors, and composition of gut microbiome. *Front Microbiol* 2020; **11:** 1402.

10 Le Bastard Q, Chapelet G, Birgand G, et al. Gut microbiome signatures of nursing home residents carrying enterobacteria producing extended-spectrum β-lactamases. *Antimicrob Resist Infect Control* 2020; **9:** 107.

11 Leo S, Lazarevic V, Girard M, et al. Metagenomic characterization of gut microbiota of carriers of extended-spectrum beta-lactamase or carbapenemase-producing Enterobacteriaceae following treatment with oral antibiotics and fecal microbiota transplantation: results from a multicenter randomized trial. *Microorganisms* 2020; **8:** e941.

12 Piewngam P, Quiñones M, Thirakittiwatthana W, Yungyuen T, Otto M, Kiratisin P. Composition of the intestinal microbiota in extended-spectrum β-lactamase-producing Enterobacteriaceae carriers and non-carriers in Thailand. *Int J Antimicrob Agents* 2019; **53:** 435–41.

13 Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host variables confound gut microbiota studies of human disease. *Nature* 2020; **587:** 448–54.

14 Verberk JDM, Vos RA, Mollema L, et al. Third national biobank for population-based seroprevalence studies in the Netherlands, including the Caribbean Netherlands. *BMC Infect Dis* 2019; **19:** 470.

15 Willems RPJ, van Dijk K, Dierikx CM, et al. Gastric acid suppression, lifestyle factors and intestinal carriage of ESBL and carbapenemase-producing Enterobacterales: a nationwide population-based study. *J Antimicrob Chemother* 2021; **77:** 237–45.

16 Kim HK, Kostidis S, Choi YH. NMR analysis of fecal samples. In: Giera M, ed. Clinical metabolomics: methods and protocols. New York, NY: Springer, 2018: 317–78.

17 Doster E, Lakin SM, Dean CJ, et al. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic Acids Res* 2020; **48:** D561–69.

18 Serafini F, Bottacini F, Viappiani A, et al. Insights into physiological and genetic mupirocin susceptibility in *Bifidobacteria*. *Appl Environ Microbiol* 2011; **77:** 3141–46.

19 El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* 2013; **11:** 497–504.

20 Annavajhala MK, Gomez-Simmonds A, Macesic N, et al. Colonising multidrug-resistant bacteria and the longitudinal evolution of the intestinal microbiome after liver transplantation. *Nat Commun* 2019; **10:** 4715.

21 Ducarmon QR, Terveer EM, Nooij S, et al. Microbiota-associated risk factors for asymptomatic gut colonisation with multi-drug-resistant organisms in a Dutch nursing home. *Genome Med* 2021; **13:** 54.

22 Araos R, Tai AK, Snyder GM, Blaser MJ, D'Agata EMC. Predominance of *Lactobacillus* spp among patients who do not acquire multidrug-resistant organisms. *Clin Infect Dis* 2016; **63:** 937–43.

23 Van Rossum T, Ferretti P, Maistrenko OM, Bork P. Diversity within species: interpreting strains in microbiomes. *Nat Rev Microbiol* 2020; **18:** 491–506.

24 Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 2016; **352:** 565–69.

25 Hu Y, Yang X, Qin J, et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 2013; **4:** 2151.

26 Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature* 2012; **486:** 2071–74.

27 Jacobson A, Lam L, Rajendram M, et al. A gut commensal-produced metabolite mediates colonisation resistance to *Salmonella* infection. *Cell Host Microbe* 2018; **24:** 296–307.e7.

28 Theriot CM, Koenigsknecht MJ, Carlson PE Jr, et al. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nat Commun* 2014; **5:** 3114.

29 Overdevest I, Haverkate M, Veenemans J, et al. Prolonged colonisation with *Escherichia coli* O25:ST131 versus other extended-spectrum beta-lactamase-producing *E coli* in a long-term care facility with high endemic level of rectal colonisation, the Netherlands, 2013 to 2014. *Euro Surveill* 2016; **21**. https://doi.org/10.2807/1560–7917.ES.2016.21.42.30376.

30 Arcilla MS, van Hattem JM, Haverkate MR, et al. Import and spread of extended-spectrum β-lactamase-producing Enterobacteriaceae by international travellers (COMBAT study): a prospective, multicentre cohort study. *Lancet Infect Dis* 2017; **17:** 78–85.