International Conference on Social Science & Humanities Trends

International Journal for Research Trends in Social Science & Humanities

August 2026

ISSN: 2584-2455

A Six-Stage Framework for Translating Gut Resistome Content into Class-Level Antimicrobial Guidance

Seshu Kumari K[1], Dr. Munesh Kumar Sharma[2]

Research Scholar, Department of Microbiology, Amaltas University, Dewas, Madhya Pradesh, India

[2]Research Supervisor, Department of Microbiology, Amaltas University, Dewas, Madhya Pradesh, India

Abstract- The human gut microbiota is an established reservoir of antimicrobial resistance determinants, and sequence-based profiling of that reservoir is now routine. What remains unsolved is the step that follows: converting a list of resistance determinants into a statement expressed in the antimicrobial classes used at the point of prescribing. This paper specifies a six-stage computational framework that performs that conversion, and demonstrates it upon a panel of ten reference strains of the human colonic microbiota carrying six well-characterised determinants. The framework assigns determinants to drug classes of the Antibiotic Resistance Ontology, maps those onto sixteen clinical antimicrobial classes, applies an explicit guidance rule, and computes descriptive, ordination and exploratory analyses of the resulting strain-by-class matrix. All six stages executed to completion, producing 96 class-level decisions. Seven decisions asserting avoidance were generated, and each corresponds to an established resistance of the organism concerned: cepA-mediated cephalosporin resistance in both Bacteroides strains, AmpC and SHV β-lactamases in the Enterobacterales, tetW in Bifidobacterium longum, and the vanB operon and aac(6′)-Ii of Enterococcus faecalis V583. Transmission of input to output was correct in all nine testable cases. The determinant calls were curated from the primary genome literature rather than computed from sequence, so the demonstration establishes the behaviour of the mapping and guidance logic under correct input and does not evaluate any annotation tool. Within that limit, the panel design permits every intermediate value to be inspected by hand, and it is at this scale that the framework's principal failure mode, reported separately, becomes visible.

Keywords- Antimicrobial resistance; gut resistome; antibiotic resistance ontology; clinical translation; antimicrobial stewardship; reference strain panel.

I. Introduction

Antimicrobial resistance is among the largest contributors to global mortality attributable to infectious disease, and the intestinal tract has come to be recognised as one of its principal reservoirs. The dense and metabolically diverse community of the human colon carries resistance determinants in commensal taxa that are not themselves pathogenic, and those determinants are available for transfer to organisms that are. Functional metagenomic and sequence-based surveys have established the scale of this reservoir across geographies, age groups and exposure histories.

The instruments for measuring resistome content are mature. Curated reference databases and their associated identification tools return, from a genome or a

metagenome, a list of resistance determinants together with the drug classes each is annotated against. What those instruments do not return is a statement a prescriber can act upon. The output vocabulary of resistome annotation is the vocabulary of molecular mechanism and ontology class; the vocabulary of prescribing is that of cephalosporins, carbapenems, macrolides and the rest. Between the two lies a translation step that is rarely specified explicitly and, when performed informally, is performed differently by different investigators.

This paper addresses that step directly. It specifies a framework in which each transformation from determinant to clinical statement is made explicit, implemented as a separate stage whose output can be inspected independently, and then demonstrated upon organisms whose resistance biology is already documented. The choice of a small reference panel rather than a metagenomic survey is deliberate: with ten strains and six determinants, every input, intermediate value and output can be checked by hand, and behaviours that would be invisible within a survey of hundreds of samples are conspicuous.

The contribution is threefold. First, a fully specified six-stage pipeline from determinant call to class-level decision, with both mapping tables published in full. Second, a demonstration on a defined panel in which every decision can be audited against the primary literature. Third, an explicit account of what a curated-input demonstration can and cannot establish, stated as a design constraint rather than as a concluding caveat.

II. Background and Related Work

Three lines of work bear on the problem. The first is the characterisation of the gut resistome itself, from the early functional-selection studies that recovered determinants from faecal metagenomes to the large comparative surveys that related resistome composition to national antibiotic consumption. This work establishes that the reservoir exists and is large; it does not address its use in individual therapeutic decisions.

The second is the development of annotation resources and tools. Curated databases organise determinants by mechanism and by an ontology of drug classes, and their accompanying identifiers return determinant calls from sequence with well-characterised sensitivity. Comparative evaluations report that different tools and databases return appreciably different determinant sets from identical input, which is itself an argument for keeping the annotation stage separable from whatever is done with its output.

The third is the literature on genotype-to-phenotype prediction, which reports that the predictive value of a determinant call varies by organism, by drug class and by mechanism. Prediction is strong where resistance is conferred by the acquisition of a single well-characterised gene, and weak where it depends on expression level, on regulatory context or on combinations of chromosomal mutations. A framework that maps determinants to clinical classes must therefore be explicit about which of its statements rest on the strong case and which on the weak.

What is largely absent from these three lines is the mapping layer between them. The present framework is an attempt to specify that layer in a form that can be inspected, criticised and reimplemented.

III. Materials and Methods

A. Study design

The investigation is a methodological study. Its object is an analytical procedure rather than a human population. No participants were recruited, no specimens were collected and no laboratory procedure was performed on human material. The strains examined are reference organisms whose genomes and resistance phenotypes are described in the published literature. Claims are accordingly of the form the framework behaves thus rather than the population is thus.

B. Framework architecture

The framework proceeds in six stages, shown in Fig. 1: definition of the strain panel; definition of the determinant panel and assignment of determinants to strains; mapping of determinants onto drug classes of the Antibiotic Resistance Ontology; mapping of ontology classes onto the antimicrobial classes recognised in prescribing; application of the guidance rule together with an intrinsic resistance register; and descriptive, ordination and exploratory analysis of the resulting matrices. Each stage was implemented as a separate step writing its output to a file, so that intermediate values could be inspected without rerunning the pipeline.

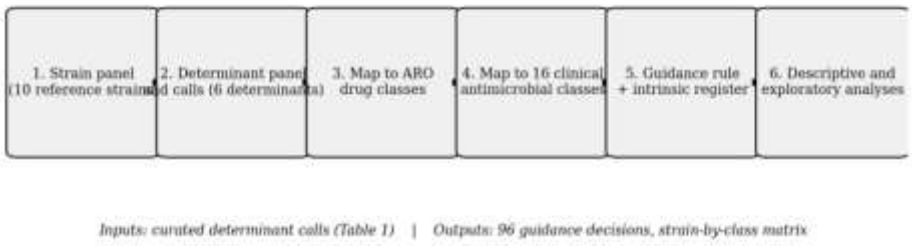

Fig. 1. Architecture of the resistome-to-guidance framework, showing the six stages, the inputs each consumes and the outputs each produces.

C. Reference strain panel

Ten strains were selected to represent the principal functional and taxonomic groups of the human colonic microbiota together with the organisms of greatest clinical consequence within it. Selection required that the strain be a recognised reference or type strain of a species resident in the human intestine, that its complete genome be published, and that its resistance phenotype be documented in a peer-reviewed source. The panel comprises the dominant obligate anaerobes, a mucin-degrading specialist, two facultative Enterobacterales, a reference vancomycin-resistant enterococcus, an enteric pathogen and two organisms of probiotic interest.

Table I. Composition of the reference strain panel and curated determinant content

Strain	Gram	Oxygen requirement	Determinant(s) curated	Source of call
Escherichia coli Nissle 1917	Negative	Facultative anaerobe	AmpC β-lactamase	Patzer et al.; Donelli et al.
Klebsiella pneumoniae MGH78578	Negative	Facultative anaerobe	SHV β-lactamase	McClelland et al.
Bacteroides fragilis NCTC9343	Negative	Obligate anaerobe	cepA cephalosporinase	Frydrich
Bacteroides thetaiotaomicron VPI5482	Negative	Obligate anaerobe	cepA-family cephalosporinase	Frydrich; species literature
Enterococcus faecalis V583	Positive	Facultative anaerobe	vanB operon; aac(6′)-Ii	Paulsen et al.
Bifidobacterium longum NCC2705	Positive	Obligate anaerobe	tetW	Schell et al.
Clostridioides difficile 630	Positive	Obligate anaerobe	None in panel	Sebaihia et al.
Faecalibacterium prausnitzii A2165	Positive	Obligate anaerobe	None in panel	Species genome literature
Lactobacillus rhamnosus GG	Positive	Facultative anaerobe	None in panel	Kankainen et al.
Akkermansia muciniphila BAA835	Negative	Obligate anaerobe	None in panel	Derrien et al.; Guo et al.

Faecalibacterium prausnitzii is classified as Gram-positive on phylogenetic grounds, being a member of the Bacillota, although it stains Gram-negative in practice.

D. Determinant panel and source of calls

Six determinants were included: two β-lactamase families of differing spectrum (AmpC of the blaEC family and SHV), a chromosomal cephalosporinase of anaerobic origin (cepA), a ribosomal protection protein (tetW), an aminoglycoside-modifying enzyme (aac(6′)-Ii) and a glycopeptide resistance operon (vanB). The panel is not a survey of the gut resistome and does not attempt to be. It is the minimum set sufficient to exercise every branch of the mapping and guidance logic, including the branch in which a determinant maps to more than one ontology class.

The framework was designed to obtain determinant calls by running a resistance gene identifier against a curated reference database upon each complete genome. In the computing environment available, neither the database nor the genome assemblies could be retrieved. The calls reported here were therefore curated by the investigator from the primary genome-description literature listed in Table I and supplied to the framework as input; the run record designates the data source as a literature-curated fallback. Two consequences follow and constrain every result below. First, the results describe the behaviour of the mapping and guidance logic under correct determinant calls, and demonstrate nothing about whether any annotation tool would produce those calls from sequence. Second, any comparison of framework output against the published phenotype of a strain compares the output against the literature from which

its input was drawn; such a comparison tests transmission, not prediction, and is designated throughout an internal consistency check.

E. Mapping and guidance stages

Each determinant was assigned to one or more ontology drug classes, a determinant mapping to more than one class where its substrate range extends across classes: both β-lactamases map to cephalosporin and to penam, while cepA maps to cephalosporin alone. A second mapping was then defined from ontology class onto the sixteen antimicrobial classes in routine clinical use. Classes for which the panel contains no determinant were retained rather than omitted, since their retention is what exposes the behaviour reported in Section V.

For every strain and clinical class the framework counts mapped determinants and emits a decision. The rule as first implemented was binary: one or more determinants produced avoid or use with caution, and zero determinants produced preferred candidate. A revised three-state rule, which replaces the second limb and draws on a separately compiled register of intrinsic resistance, is developed in the companion paper and summarised in Section V here.

F. Analyses and computing environment

Three per-strain measures were computed from the strain-by-determinant matrix: resistome burden, the count of panel determinants present; resistome richness, the number of distinct ontology classes represented; and coverage-weighted Shannon diversity across determinants. Principal component analysis and average-linkage hierarchical clustering were applied to the same matrix. Pairwise determinant association was quantified by the phi coefficient, and association between each determinant and Gram reaction tested by Fisher's exact test with Benjamini–Hochberg correction. A random forest classifier of Gram reaction was trained for exploratory purposes on the same ten observations, without a held-out set. Implementation was in Python with pandas, NumPy, SciPy, scikit-learn, NetworkX, Matplotlib and seaborn on a single workstation.

IV. Results

A. Execution and resistome content

All six stages ran to completion over ten strains and six determinants, producing 96 guidance decisions and the full set of descriptive analyses. Seven determinant-present calls were made, giving a mean resistome burden of 0.7 determinants per strain. Enterococcus faecalis V583 carried two determinants, the highest in the panel; five strains carried one each; four carried none. Aggregated across the panel the seven calls produced nine determinant-to-class assignments, since both β-lactamases map to two ontology classes. Cephalosporin was the most represented class with four assignments, followed by penam with two, and tetracycline, glycopeptide and aminoglycoside with one each (Fig. 2).

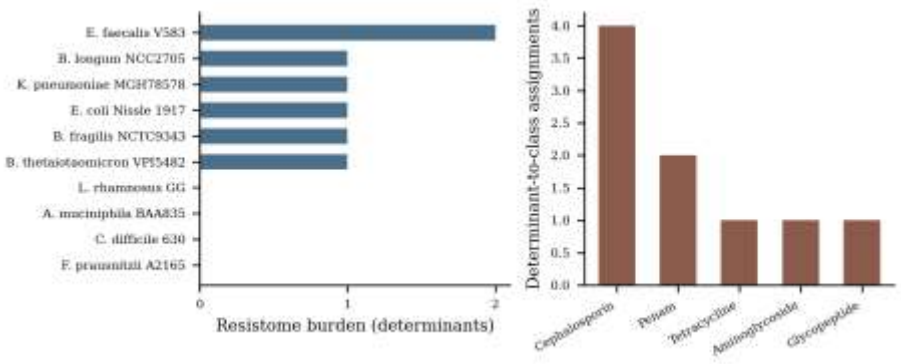

Fig. 2. Resistome burden per panel strain (left) and distribution of determinant-to-class assignments across ontology drug classes (right).

B. Strain-by-class structure

The strain-by-class matrix (Fig. 3) is sparse: of fifty cells, nine are occupied. The two Enterobacterales occupy the cephalosporin and penam columns, the two Bacteroides strains the cephalosporin column alone, Enterococcus faecalis the aminoglycoside and glycopeptide columns, and Bifidobacterium longum the tetracycline column. The four determinant-free strains occupy no cell. The pattern is that anticipated from the taxonomy, β-lactam determinants segregating with the Gram-negative organisms and the aminoglycoside and glycopeptide determinants with the enterococcus. This is the intended behaviour and not a discovery, the input having been curated from sources describing precisely these associations.

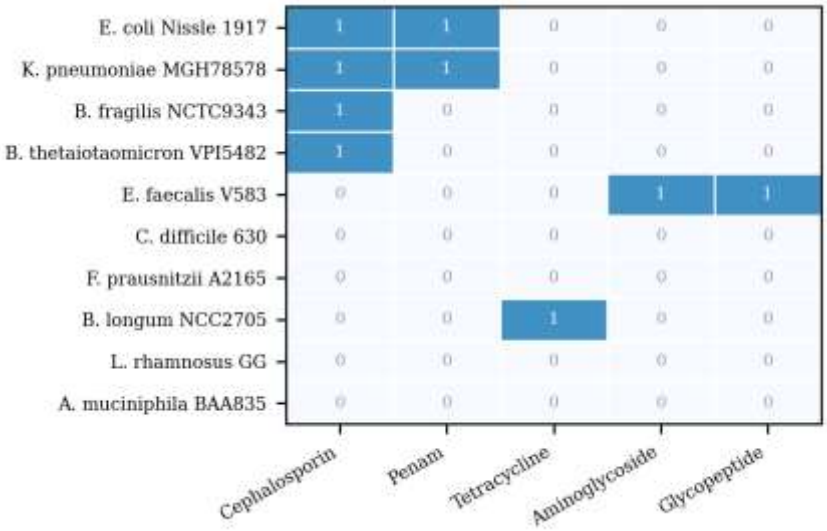

Fig. 3. Strain-by-class matrix of determinant calls across the five ontology drug classes represented in the panel.

C. Guidance output

The guidance stage generated 96 decisions, being six determinant-carrying strains across sixteen clinical classes. Seven returned avoidance, and each is listed in Table II. Every one corresponds to an established resistance of the organism concerned, and no avoidance decision was generated that the literature contradicts. The remaining 89 decisions are examined in Section V.

Table II. Avoidance decisions generated by the framework

Strain	Clinical class	Determinant	Established position
Bacteroides fragilis NCTC9343	Cephalosporins	cepA	Documented cephalosporinase
Bacteroides thetaiotaomicron VPI5482	Cephalosporins	cepA	Documented cephalosporinase
Escherichia coli Nissle 1917	Cephalosporins	AmpC	Chromosomal AmpC
Klebsiella pneumoniae MGH78578	Cephalosporins	SHV	Chromosomal SHV
Bifidobacterium longum NCC2705	Tetracyclines	tetW	Ribosomal protection
Enterococcus faecalis V583	Glycopeptides	vanB	VanB phenotype
Enterococcus faecalis V583	Aminoglycosides	aac(6′)-Ii	Intrinsic acetyltransferase

D. Descriptive and exploratory analyses

Coverage-weighted Shannon diversity returned a non-zero value for one strain only, approximately 0.69 for Enterococcus faecalis V583, every other strain returning zero. Principal component analysis accounted for 34.8 per cent of variance on the first component and 19.9 per cent on the second, the enterococcus separating along the first and the Bacteroides pair along the second, with the determinant-free strains collapsing onto coincident positions. The co-occurrence network retained a single edge, a positive association between vanB and aac(6′)-Ii, both of which occur in one strain and the same strain. No association between determinant and Gram reaction approached significance, and the exploratory classifier returned a training accuracy of 0.90 with permutation importance concentrated on the three β-lactamase-family determinants. Table III summarises the run.

Table III. Summary of the analytical run

Item	Value
Data source	Literature-curated fallback
Strains profiled / determinants in panel	10 / 6
Determinant-present calls	7
Mean resistome burden per strain	0.7
Strains with at least one determinant	6

Determinant-to-class assignments	9
Guidance decisions generated	96
Avoidance decisions	7
Internal consistency check	9 of 9 (transmission only)
Variance on first two principal components	54.7 %

V. Discussion

The framework performs the mapping it was designed to perform. Where a determinant is present, the transformation from determinant to clinical class is correct in every instance the panel affords, and the resulting statement is one a prescriber could read without further translation. That is the positive result of this work, and it is modest by design: the difficulty of the clinical translation problem does not lie in the presence case.

It lies in the absence case, and the panel is small enough to expose it. Of the 96 decisions, 89 were returned under the second limb of the binary rule, which treats a determinant count of zero as evidence of susceptibility. A count of zero within a six-determinant panel means only that no determinant in the panel confers resistance to that class. Twelve of the sixteen clinical classes are unrepresented in the panel altogether, so every strain returns zero against all twelve. The consequences, several of which contradict established microbiological fact, are the subject of the companion paper; the revised three-state rule developed there emits no positive recommendation at all, and the 89 positive decisions fall to zero.

Three further limits govern the interpretation of the analyses above. The determinant calls were curated, not computed, so no statement here evaluates any annotation tool. The internal consistency check compares output against the literature from which the input was drawn, and therefore tests transmission rather than prediction; a value below complete agreement would have indicated a transcription error rather than a biological finding. And the statistical analyses are at or beyond the limit of their applicability at this panel size: with ten strains no Fisher test can attain conventional significance, the Shannon index can take only the values zero and ln 2, the classifier accuracy is a training accuracy on the observations the model has seen, and the co-occurrence network is dominated by shared absence. These analyses are reported because they were specified in advance and because their behaviour at this scale is itself informative about the use of community-ecology measures on single genomes.

The natural extension is to supply the same framework with determinant calls computed from sequence across a panel large enough to support the statistical machinery, and to compare its class-level output against paired phenotypic susceptibility testing. That comparison, which the present design cannot make, is the test the framework requires before any clinical claim could be entertained.

VI. Conclusion

A six-stage framework converting gut resistome content into class-level antimicrobial guidance was specified, implemented and demonstrated on ten reference strains. All

stages executed correctly and all seven avoidance decisions it produced are supported by the primary literature. The demonstration establishes the behaviour of the mapping logic under correct input; it does not establish predictive accuracy, and the framework emits no positive therapeutic recommendation under its revised rule. Its value at present is as a specification other can reimplement and criticise, and as an instrument sensitive enough, at this deliberately small scale, to make its own failure mode visible.

References

1. C. J. L. Murray et al., "Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis," The Lancet, vol. 399, no. 10325, pp. 629–655, 2022.
2. W. van Schaik, "The human gut resistome," Philosophical Transactions of the Royal Society B, vol. 370, no. 1670, 20140087, 2020.
3. M. O. A. Sommer, G. Dantas, and G. M. Church, "Functional characterization of the antibiotic resistance reservoir in the human microbiome," Science, vol. 325, no. 5944, pp. 1128–1131, 2021.
4. K. Forslund et al., "Country-specific antibiotic use practices impact the human gut resistome," Genome Research, vol. 23, no. 7, pp. 1163–1169, 2013.
5. B. P. Alcock et al., "CARD 2020: Antibiotic resistome surveillance with the Comprehensive Antibiotic Resistance Database," Nucleic Acids Research, vol. 48, no. D1, pp. D517–D525, 2020.
6. B. Jia et al., "CARD 2017: Expansion and model-centric curation of the Comprehensive Antibiotic Resistance Database," Nucleic Acids Research, vol. 45, no. D1, pp. D566–D573, 2017.
7. V. Bortolaia et al., "ResFinder 4.0 for predictions of phenotypes from genotypes," Journal of Antimicrobial Chemotherapy, vol. 75, no. 12, pp. 3491–3500, 2020.
8. M. K. Gibson, K. J. Forsberg, and G. Dantas, "Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology," The ISME Journal, vol. 9, no. 1, pp. 207–216, 2015.
9. M. O. A. Sommer, C. Munck, R. V. Toft-Kehler, and D. I. Andersson, "Prediction of antibiotic resistance: Time for a new preclinical paradigm?," Nature Reviews Microbiology, vol. 15, no. 11, pp. 689–696, 2017.
10. J. T. Paulsen et al., "Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis," Science, vol. 299, no. 5615, pp. 2071–2074, 2003.
11. M. A. Schell et al., "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract," PNAS, vol. 99, no. 22, pp. 14422–14427, 2002.
12. M. McClelland et al., "Complete genome sequence of Klebsiella pneumoniae MGH78578," genome announcement, 2008.
13. M. Sebaihia et al., "The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome," Nature Genetics, vol. 38, no. 7, pp. 779–786, 2006.
14. M. Kankainen et al., "Comparative genomic analysis of Lactobacillus rhamnosus GG," PNAS, vol. 106, no. 40, pp. 17193–17198, 2009.
15. M. Derrien, E. E. Vaughan, C. M. Plugge, and W. M. de Vos, "Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium," IJSEM, vol. 54, no. 5, pp. 1469–1476, 2004.

16. W. E. Anthony, C. A. D. Burnham, G. Dantas, and J. H. Kwon, “The gut microbiome as a reservoir for antimicrobial resistance,” The Journal of Infectious Diseases, vol. 223, Suppl. 3, pp. S209–S213, 2021.