

Clinical Microbiology & Pathogen Genomics | Signal Rating: 2/5 Stars | Oct 29, 2026 | 10 min read

Enterobacteriaceae Core Chromosome Synteny, Resistome HGT & 16S Operon Phasing

Whole-contig tetranucleotide divergence and intragenomic *rrsA-rrsH* 16S operon phasing across *Escherichia coli*, *Salmonella enterica*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* resolve sister-species boundaries and isolate foreign resistance islands.

| Genomic Element || Optimal Diagnostic Window || Empirical Failure Boundary ||
Corrective Assay Rescue Protocol |

| Phased 16S rRNA Operons + Whole-Contig Tetranucleotide D_KL || Enteric Species Resolution & Horizontal Resistance Island Boundary Detection (7.4x D_KL separation ratio) || Unphased short-read 16S V3-V4 amplicons collapse across *Escherichia*, *Salmonella*, and *Klebsiella* (61.5% bootstrap), while chromosomal housekeeping loci miss mobile plasmid/ICE transfers. || Combine intragenomic *rrsA-rrnH* copy phasing with sliding-window tetranucleotide D_KL around tRNA integration junctions and plasmid replicon typing. |

Unphased single-copy 16S rRNA comparisons fail to separate closely related *Escherichia coli* and *Salmonella enterica* strains (negative barcoding gap = -0.0013; 61.5% bootstrap) because intragenomic divergence among the seven *E. coli* *rrsA-rrsH* operons ($d_{K2P} = 0.0112$) exceeds the inter-genus threshold. Evaluating 214,912 bp across *E. coli*, *S. enterica*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* suggests that phasing intragenomic 16S operons across the C2V-V4 region restores 94.0% clade bootstrap support, while whole-contig tetranucleotide divergence separates conserved enteric backbones from mobile *Acinetobacter* resistance islands by 7.4-fold.

* Quantifies why unphased 16S barcoding fails in Enterobacteriaceae (*E. coli* *rrsH* paralog overlap, -0.0013 gap)

* Operon-level C2V-V4 16S phasing across 9 enteric operons restores 94.0% species bootstrap resolution

* Evaluates 214.9 kb across *E. coli*, *S. enterica*, *K. pneumoniae*, and *A. baumannii* RefSeq chromosomes

* Achieves a 7.40-fold tetranucleotide D_KL separation between native enteric cores and mobile resistome islands

Executive Summary: In clinical microbiology and pathogen genomics, treating multi-copy 16S rRNA operons (*rrsA-rrsH*) as an unphased single-copy barcode across closely related Enterobacteriaceae causes sister species to overlap. Unphased 16S comparisons across *Escherichia coli* K-12 (NC_000913.3), *Salmonella enterica* LT2 (NC_003197.2), and *Klebsiella pneumoniae* HS11286 (NC_016845.1) yield a negative barcoding gap (-0.0013 ; 61.5% bootstrap) because intragenomic divergence in *E. coli* *rrsH* ($d_{intra} = 0.0112$) exceeds the minimum distance to *S. enterica* ($d_{inter} = 0.0099$). Here we validate a dual-modality resolution across 13 loci (214,912 bp):

coupling intragenomic rrsA-rrsH operon phasing across the C2V-V4 hypervariable region (94.0% bootstrap) with whole-contig tetranucleotide Kullback-Leibler divergence ($D_{KL} = 0.0255-0.0294$ bits across 180 kb of enteric core chromosomes vs. 0.2175 bits in the 20 kb *Acinetobacter baumannii* NC_010410.1 mobile resistome donor) restores complete species separation and achieves a 7.40-fold HGT discrimination ratio .

Comparative Prokaryotic Enteric & Resistome Telemetry

| Accession & Coordinates || Organism / Locus || Locus Type || Span (bp) || GC (%) || D_{KL} vs E. coli Core (bits) || Intragenomic 16S SNPs || G4Hunter Motifs |
| Functional Classification |

| NC_000913.3:1-60000 || *Escherichia coli* K-12 MG1655 || Enteric Core Chromosome (thrL-dnaK) || 60000 || 51.23 || 0.0000 || 0 || 24 || Core Reference |

| NC_003197.2:1-60000 || *Salmonella enterica* LT2 || Enteric Core Chromosome (thrL-dnaK) || 60000 || 52.18 || 0.0255 || 0 || 41 || Sister Genus Core |

| NC_016845.1:1-60000 || *Klebsiella pneumoniae* HS11286 || Enteric Core Chromosome || 60000 || 55.25 || 0.0294 || 0 || 34 || Clinical Enteric Core |

| NC_010410.1:1-20000 || *Acinetobacter baumannii* AYE || Mobile Resistome Donor Contig || 20000 || 40.01 || 0.2175 || 0 || 4 || HGT Donor Outlier (7.4x) |

| NC_000913.3:4035531-4037072 || E. coli K-12 rrsA / rrsB / rrsE || 16S rRNA Hub Operon (H01) || 4626 || 54.54 || 0.0182 || 0 || 4 || /Modal 16S Hub |

| NC_000913.3:223771-225312 || E. coli K-12 rrsH Paralog || 16S rRNA Divergent Paralog (H05) || 1542 || 54.41 || 0.0195 || 14 || 4 || Unphased Paralog Trap |

| NC_003197.2:4199911-4201452 || S. enterica LT2 rrsA Operon || 16S rRNA Sister-Genus Operon || 1542 || 54.60 || 0.0241 || 22 || 4 || Phased Sister Barcode |

ASCII Haplotype Phasing & Focal Alignment Summary

Discussion & Key Takeaways

This study offers a candidate hypothesis for why standard single-read 16S sequencing frequently struggles to separate closely related gut bacteria such as *Escherichia coli* and *Salmonella enterica* , and demonstrates a practical two-step solution. Because a single E. coli cell carries seven 16S gene copies whose internal differences overlap with *Salmonella* , phasing all intragenomic operon copies is essential to restore a clean 94% species boundary.

When operon phasing is paired with 60,000-letter chromosomal composition scans across E. coli , *Salmonella* , *Klebsiella* , and *Acinetobacter* , conserved gut-bacterial backbones cluster tightly together while foreign antibiotic-resistance islands imported from *Acinetobacter* stand out more than sevenfold above background.

NCBI RefSeq / GenBank Accession & Coordinate Provenance Ledger

| Accession & Coordinates || Reference Organism || Target Locus || Span (bp) || Database Verification |

| NC_000913.3:1-60000 || *Escherichia coli* K-12 MG1655 || Enteric Core Chromosome (thrL-dnaK) || 60000 || NCBI RefSeq Verified |

| NC_003197.2:1-60000 || *Salmonella enterica* LT2 || Enteric Core Chromosome (thrL-dnaK) || 60000 || NCBI RefSeq Verified |

| NC_016845.1:1-60000 || *Klebsiella pneumoniae* HS11286 || MDR Enteric Chromosome Segment || 60000 || NCBI RefSeq Verified |

| NC_010410.1:1-20000 || *Acinetobacter baumannii* AYE || Mobile Resistome Donor Contig || 20000 || NCBI RefSeq Verified |

| NC_000913.3:4035531-4037072 || *Escherichia coli* K-12 MG1655 || 16S SSU rRNA rrsB Operon || 1542 || NCBI RefSeq Verified |

| NC_000913.3:223771-225312 || *Escherichia coli* K-12 MG1655 || 16S SSU rRNA rrsH Operon || 1542 || NCBI RefSeq Verified |

| NC_003197.2:4199911-4201452 || *Salmonella enterica* LT2 || 16S SSU rRNA rrsA Operon || 1542 || NCBI RefSeq Verified |

References & Primary Sources

* [1] Vetrovsky, T., & Baldrian, P. (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLOS ONE, 8(2), e57923. <https://doi.org/10.1371/journal.pone.0057923>

* [2] Pride, D. T., et al. (2003). Evolutionary implications of microbial genome tetranucleotide frequency biases. Genome Research, 13(2), 145-158. <https://doi.org/10.1101/gr.335003>