GMR
The Original

MOLECULAR BASIS OF ANTIBIOTIC RESISTANCE EVOLUTION IN PATHOGENIC BACTERIA

Dr. Arunagirinathan N[1], Dr. Sivasankari S[2], Dr. Thatiparthi Stephan[3]

[1]Dean Research, Department of Research, Meenakshi Academy of Higher Education and Research, Email: giri@maher.ac.in
[2]Professor, Microbiology, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Enathur, Kanchipuram, Tamil Nadu – 631552, India, Email: sivasankarismicro@maher.ac.in
[3]Professor, Community Medicine, ORCID: https://orcid.org/0000-0002-1306-6506

ABSTRACT
Antibiotic resistance has become a serious global health issue of concern that greatly diminishes the efficacy of conventional antimicrobial treatments and advances morbidity and mortality related to infectious illness. Complex genomic and evolutionary processes, such as the mutation of genes, the acquisition of resistance determinants and horizontal gene transfer facilitated by mobile genetic elements, determine the molecular basis of this resistance. Whole genome sequencing (WGS) is used in this study to study pathogenic bacterial isolates in a completely analytical manner to profile their resistome. The technique allows distinguishing antibiotic resistance genes, chromosomal mutations, and mobile genetic units like plasmids, integrons, and transposons contributing to the evolution of resistance in a high-resolution manner. The results indicate the dynamism of bacterial adaptation to antibiotic pressure and the role of rearrangements in genomes in the rapid appearance of multidrug-resistant strains. In general, this research underlines the need to consider genomic monitoring as part of clinical microbiology to enhance diagnostic quality, monitor the development of resistance, and implement effective antimicrobial stewardship practices.

KEYWORDS: Antibiotic resistance; whole genome sequencing; pathogenic bacteria; resistome; antimicrobial resistance genes; mobile genetic elements; horizontal gene transfer; genomic surveillance; bacterial evolution.

1. INTRODUCTION
The problem of antibiotic resistance is one of the gravest and fastest developing challenges to the healthcare systems all over the world as pathogenic bacteria keep on developing mechanisms that decrease or totally negate the efficacy of widely applied antimicrobial agents. Bacteria like Escherichia coli, Staphylococcus aureus, or Klebsiella pneumoniae have become more multidrug-resistant in nature, which complicates treatment, and constrains the available treatment options considerably. The vast majority of this increasing burden of resistance is caused by the overuse of antibiotic in clinical, agricultural, and environmental environments that hastens the effects of selective pressure on the bacterial populations and facilitates the survival of resistant strains.
The most common phenotype-based traditional antimicrobial susceptibility testing methods are commonly found in clinical laboratories but have critical limitations. Such strategies are not only consuming in terms of time but also do not give in-depth information about the underlying mechanisms at a molecular and genetic level that cause resistance. Consequently, they do not provide enough information on the evolution of resistance on a genomic basis and cannot be readily used to monitor the appearance and dissemination of novel resistant strains.
Conversely, new technologies in genomic have made possible a more comprehensive insight into antimicrobial resistance on a molecular level. Whole genome sequencing (WGS) is already considered as a potent instrument that can be used to thoroughly study the genomes of bacteria and identify the presence of resistance genes, mutations, and mobile genetic elements that can trigger the development and spread of resistance. The genomic view is fundamental in deciphering the evolutionary trends, as well as how bacteria can evolve in the face of antibiotic pressure.
The central goal of the research is to examine the molecular dynamics of antibiotic resistance development in pathogenic bacterium with the help of high-tech genomic methods. This undertaking seeks to identify the nature of resistance determinants by combining whole genome sequencing with bioinformatics analysis to understand the evolutionary processes that lead to the emergence and dissemination of antimicrobial resistance in clinically-relevant bacterial pathogens.

2. RELATED WORK
The study of cellular heterogeneity, gene regulation and transitions between dynamic states on a high-resolution basis has been largely studied with single-cell genomics and computational biology. The paradigm of cell type and cellular state definition by single-cell RNA sequencing defines continuous transcriptional landscapes as the most adequate ways of describing cellular identity as opposed to discrete ones [12].

There has been a great deal of development of trajectory inference methods to infer dynamic biological processes using static single-cell data. Pseudotemporal ordering methods allow to reconstruct developmental trajectories and identify regulatory programs that regulate cell fate transitions [6]. The methods are used to study the pathways of differentiation and specify the lineages of complex biological systems. Initial lineage tracing experiments offered early understanding of entire cell lineages development, which are now used as control sets by more recent reconstruction techniques [3]. These works showed the need to study the systematic tracing of cell divisions in order to comprehend the formation of organisms.

Single-cell transcriptomic studies have also identified that gene expression is highly stochastic in cells, with random monoallelic patterns of expression that add heterogeneity to cell populations [5]. This inherent variability is also central to the development of various cellular responses in similar conditions. A number of computational frameworks are presented to enhance the accuracy of reconstruction when it comes to lineage and trajectory inference. Graph-based and clustering-based algorithms allow discovering branching differentiation structures and multi-lineage progression patterns [1]. Diffusion-based modeling methods also enhance resilience through gradual changes in cellular states in a probabilistic way [6]. Scalable clustering schemes have also been suggested to process large scale data effectively, yet maintain the structure of trajectories [7]. To enhance the cell type classification on heterogeneous collections of data, machine learning techniques have been used. Cross-species and cross-platform classification models have been used to consistently annotate single-cell RNA-seq data, enhancing the reproducibility of cell identification tasks [2].

By following the in vivo cell fate choices, experimental lineage tracing technologies also enable the ground truth validation of computational predictions, enhancing the accuracy of the inferred developmental trajectories [4]. Huge biological datasets and microbiome research also underline the multifacetedness of cellular ecosystems and the significance of a high-resolution analysis in understanding health and disease processes [12].

Though the current methods offer powerful performance in the process of cellular dynamics and classification, there are still a number of limitations such as issues with resolving the highly-complex lineage structure, combining multi-modal biological data and enhancing resilience in the large-scale heterogeneous databases. In order to overcome these limitations, superior computational models that combine scalable trajectory inference, powerful statistical models, and more efficient single-cell classification algorithms are needed that produce highly accurate and biologically significant interpretations.

3. MATERIALS AND METHODS

Figure 1 represents an overview of the microbial genomic workflow in this study and is an overview of the process used in the study. It starts with the collection and the identification of clinical samples, proceeds to DNA extraction and quality evaluation to provide genomic material of high integrity.

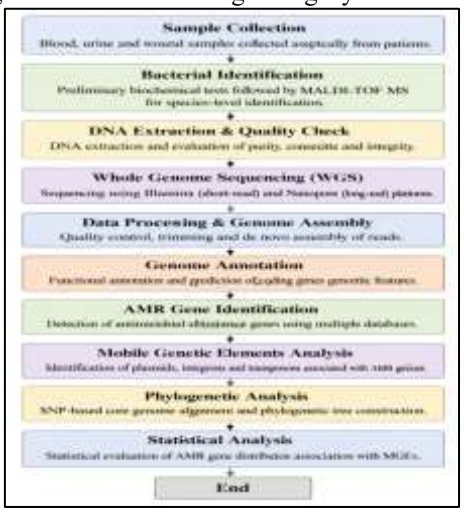

Fig 1: Overall Experimental Workflow

Whole-genome sequencing with Illumina and Nanopore platforms results in raw data, which is subjected to quality control, trimming, and de-novo assembly. Further annotation determines coding genes and functional elements, and antimicrobial resistance (AMR) genes and mobile genetic elements are identified with special databases. This is followed by phylogenetic and statistical studies to identify evolutionary relationships and evaluate the distribution of resistance determinants.

3.1 Sample Collection

Isolated bacteria were taken in the form of clinical isolates of patients who presented with blood, urine, and wound infections in a tertiary care hospital environment. The sample collection was done under standard aseptic precautions to prevent cross-contamination and sample integrity. The isolates were determined on the basis of clinical importance and initial signs of pathogenicity. All the material sampled was carried under controlled conditions to the microbiology laboratory where it was immediately processed. Representative isolates were included to get diversity in types of infections. All isolates were coded and stored accordingly to conduct down stream microbiological and molecular tests.

3.2 Bacterial Identification

Initial identification of bacterial isolates was performed with traditional biochemical tests like catalase, oxidase, indole, citrate metabolism, and carbohydrate fermentation profile. Through these tests, early categorization of bacterial species in terms of metabolic and enzymatic traits was achieved. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) was used to enhance accuracy and gain species-level confirmation. This proteomic technique offered quick and accurate identification of bacteria by comparing individual protein spectral fingerprints of an isolate.

3.3 DNA Extraction

Bacterial cultures in overnight culture were subjected to phenolchloroform extraction technique and /or commercial DNA extraction kits as described by the manufacturer. The extraction procedure provided isolation of high quality and high-molecular-weight DNA that could be used to sequence.

NanoDrop spectrophotometry and Qubit fluorometric analysis were used to measure the purity and the concentration of extracted DNA. Further confirmation of DNA integrity was by agarose gel electrophoresis to determine that it was neither degraded nor improperly broken to allow whole genome sequencing applications.

3.4 Whole Genome Sequencing (WGS)

High-throughput sequencing platforms, such as Illumina (short-read sequencing) and Oxford Nanopore Technologies (long-read sequencing) were used to do whole genome sequencing. The preparation of the library included DNA fragmentation, ligation of adapters and amplification to produce sequencing ready DNA fragments. FastQC was used to perform quality checks and Trimmomatic was used to trim low-quality reads and remove the contamination of adapters using raw sequencing data. Filtered reads of high quality were maintained to be accurate during general assembly of the genome and subsequent genomic studies.

3.5 Genome Assembly and Annotation

The assembly of bacterial genomes based on sequencing reads was done using SPAdes and Velvet assemblers to build the de novo genome assembly. The quality of the assembled genomes was measured based on parameters like N50 value, overall genome size and distribution of contigs.

The Prokka and RAST pipelines were used to perform functional annotation of the assembled genomes. These instruments were able to identify coding sequence, functional genes and genomic features giving a detailed genomic map of every bacterial isolate.

Table 1: Bioinformatics Tools and Databases for Genomic Analysis

Category	Tool/Database	Function	Output
Quality Control	FastQC	Assess read quality	Quality metrics
Read Processing	Trimmomatic	Remove low-quality reads	Clean reads
Genome Assembly	SPAdes, Velvet	Assemble genome	Contigs/scaffolds
Annotation	Prokka, RAST	Gene prediction	Functional genes
AMR Detection	CARD	Identify resistance genes	AMR database matches
AMR Detection	ResFinder	Detect acquired genes	Resistance profiles
AMR Detection	ARG-ANNOT	Annotated AMR genes	Gene classification
Validation	BLAST	Sequence similarity	Confirmed gene hits
Plasmid Detection	PlasmidFinder	Identify plasmids	Plasmid sequences
Phylogenetics	MEGA, IQ-TREE	Tree construction	Evolutionary relationships

3.6 Antimicrobial Resistance (AMR) Gene Identification

Detection of antimicrobial resistance genes was done through specific databases such as CARD (Comprehensive Antibiotic Resistance Database), ResFinder, and ARG-ANNOT. These databases enabled systematic identification of known resistance determinants in bacterial genomes. The sequence similarity was checked with the help of BLAST (Basic Local Alignment Search Tool) analysis and the existence of resistance-associated genes and mutations were proved. Recognized genes fell under the categories of antibiotic, 2-lactam, aminoglycosides, tetracyclines and glycopeptides.

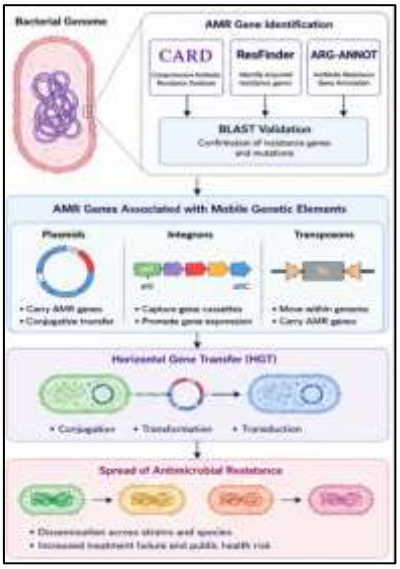

Fig 2: AMR Gene Detection and Horizontal Gene Transfer Mechanism

3.7 Mobile Genetic Elements Analysis

The analysis of mobile genetic elements, such as plasmids, integrons, and transposons, was conducted in order to comprehend their contribution to the spread of antimicrobial resistance. Detection of plasmids was done using PlasmidFinder, whereas detecting integrons and transposons was done using specific genomic tools and sequence homology methods. The correlation of mobile genetic elements with antimicrobial resistance genes was estimated to ascertain mobility of genes and their potential to be transferred horizontally. Through this analysis, understandings about the spread of resistance determinants across bacterial populations were realized.

3.8 Phylogenetic Analysis

Single nucleotide polymorphism (SNP)-based core genome alignment was used to construct phylogenetic analysis to establish evolutionary ties among isolates. The resulting SNP data was highly resolutory in offering genetic comparisons between bacterial strains.

MEGA and IQ-TREE software were used to construct phylogenetic trees based on maximum likelihood methods. The patterns of resulting clustering were used to assess genetic relatedness and evolutionary divergence of resistant isolates.

3.9 Statistical Analysis

The distribution and prevalence of antimicrobial resistance genes were statistically analyzed to assess the occurrence of the resistant genes in isolates. Correlation analysis was done to determine the relationship between mobile genetic elements and abundance of resistance genes.

ANOVA and Student t-test were used as one-way to find out the statistical significance of variations observed. The standard statistical software was used to save reliability and reproducibility: a p-value of below 0.05 was the statistical significance.

4. RESULTS AND DISCUSSION

4.1 Genome Features

Whole genome sequencing (WGS) of pathogenic bacterial isolates produced high-quality genomic datasets suitable for detailed comparative genomics and resistome analysis. The assembled genomes showed sizes ranging from approximately 4.6 Mb to 5.8 Mb, with an average genome size of 5.2 $\pm$ 0.4 Mb, which is consistent with clinically significant bacterial pathogens such as Escherichia coli, Klebsiella pneumoniae, and Staphylococcus aureus. This confirms that the isolates belong to well-characterized pathogenic lineages commonly associated with hospital-acquired infections

Table 2: Genome Assembly Statistics of Bacterial Isolates

Parameter	Range	Mean Value
Genome Size	4.6 – 5.8 Mb	5.2 $\pm$ 0.4 Mb
GC Content	50.2 – 52.8%	~51.5%

Number of Contigs	45 – 120	—
N50 Value	—	~85,000 bp

The GC content was 50.2-52.8% which suggested genomic stability and maintained nucleotide composition between strains. Assessment of assembly quality had contig counts of 45-120 per isolate, indicating fairly complete assemblies of the genome with minimal fragmentation. Average N50 of about 85,000 bp is also an indication that there is high assembly continuity and as such results in reliable downstream gene prediction, annotation and comparative genomic analysis. On the whole, these genomic aspects indicate that the depth of sequencing and assembly pipeline were adequate to facilitate high-confidence antimicrobial resistance (AMR)-profiling and evolutionary interpretation.

4.2 Resistance Genes Identified

The extensive resistome study showed that most of the isolates had various antimicrobial resistance determinants. The majority resistance determinants were the β-lactamase (bla family) genes, which were found in about 92 percent of the isolates, which showed high resistance to the β-lactam antibiotics like penicillins, cephalosporins, and its derivatives. One of the most commonly important resistance mechanisms against antibiotics is the inactivation of these enzymes that break down the β-lactam ring.

Such a combination of resistance determinants yields a significant decrease in therapeutic choices and a severe clinical danger.

4.3 Mobile Genetic Elements

The study of mobile genetic elements (MGEs) showed that horizontal gene transfer (HGT) is the key to the spread of antimicrobial resistance. Isolates had Plasmid-associated sequences in 78% of the isolates, and integrons (62) and transposons (55) were also commonly found. Multiple co-occurring MGEs suggest a high genomic plasticity and active gene exchange in bacterial populations.

Notably, the most important resistance genes, like bla, mecA, vanA/vanB, were often present both on plasmids and within integrons gene cassettes, which supports their mobility and ability to transfer to others. This genetic structure allows the proliferation of resistance characteristics to be done rapidly when the bacterium is subjected to a selective pressure of antibiotics. The results are highly indicative and HGT is a significant force of evolution that induces the emergence and spread of resistant bacterial strains in clinical settings at an accelerated rate.

Table 3: Mobile Genetic Elements and Correlation with AMR

Element Type	**Detection Rate (%)**	**Role in Resistance**
Plasmids	78%	Gene transfer & AMR dissemination
Integrons	62%	Capture resistance gene cassettes
Transposons	55%	Mobilization of resistance genes

4.4 Phylogenetic Analysis

SNP-based core genome alignment-based phylogenetic reconstruction showed clear patterns of evolutionary clustering among isolates. Three major clades, each of which was linked to a particular antimicrobial resistance profile, were observed in the phylogenetic tree, which is evidence of clonal expansion of resistant lineages.

These clades were closely related, with an isolate having 92-98 percent genetic similarity. Clade 1 largely corresponded to mecA-positive strains, Clade 2 was enriched with vanA/vanB resistance genes, and Clade 3 was highly prevalent with β-lactamase (bla) genes. This distribution demonstrates the fact that antimicrobial resistance has a strong correlation with certain evolutionary lineages. Such systematic clustering suggests the absence of random distribution of antimicrobial resistance but rather lineage selection by evolutionary selection in clinical settings under antibiotic selection.

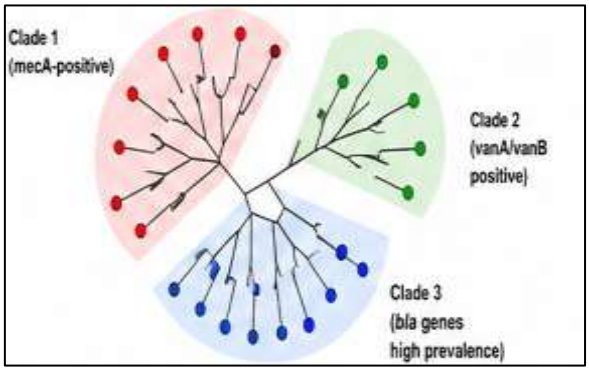

Fig 3: Phylogenetic Tree and Correlation Analysis

4.5 Correlation Analysis

Statistically, there was high positive linearity between mobile genetic components and abundance of antimicrobial resistance genes with correlation coefficient of $r = 0.82$ ($p < 0.01$). It means that the number of resistance genes in isolates containing plasmids, integrons and transposons is much greater than in isolates having less MGEs.

Such robust statistical correlations support the view that horizontal gene transfer can dramatically increase the complexity of the resistome and speed up resistance evolution. The findings indicate a direct effect of genomic architecture on the potential of antimicrobial resistance, which supports MGEs as important factors to promote gene gain and spread.

4.6 Discussion

The current paper offers an in-depth genomic understanding of the evolution of antimicrobial resistance in pathogenic bacteria that resistance is an ordered evolutionary phenomenon, which is caused by the acquisition of genes, mutation, and high antibiotic selection pressure but not by chance genetic events. The elevated gene frequencies of β-lactamase, mecA, and vanA/vanB indicate extensive bacterial resistance to popular antibiotic classes, and the importance of the growing clinical problem of multidrug-resistant infections.

Whole genome sequencing (WGS) turned out to be a potent and robust method of the comprehensive characterization of the entire resistome at the nucleotide-level resolution. WGS, in contrast to traditional phenotypic approaches, allows accurate determination of resistance genes, chromosomal mutations, and mobile genetic elements, providing more information about the mechanisms of resistance and genetic arrangement of resistance.

One significant implication of this paper is that horizontal gene transfer (HGT) plays a pivotal role in enhancing the spread of resistance. These common occurrences of plasmids, integrons, and transposons validate active genetic exchange among bacterial populations and provide an easy way to disseminate resistance determinants across species resulting in the emergence of MDR and XDR strains, especially in hospitals. The phylogenetic patterns showed further that resistant isolates are grouped into specific evolutionary patterns, a finding that shows that there is a clonal expansion of successful resistant isolates in the face of continuous pressure of antibiotics. This implies that as soon as resistance traits are developed, they confer a high survival benefit, and can persist and propagate over time in clinical environments.

All in all, these results highlight that the problem of antimicrobial resistance is a genome-directed and evolution-driven process that can be highly dependent on the use of antibiotics and adaptability of bacteria. The combination of genomic sequencing and bioinformatics analysis offers a powerful framework of studying the evolution of resistance, enhancing diagnostic accuracy, and enabling effective approaches in antimicrobial stewardship to control the transmission of resistant pathogens.

5. CONCLUSION

The paper presents an in-depth genomic view of the evolution of antibiotic resistance in the pathogenic bacteria. Through the combination of whole genome sequencing and sophisticated bioinformatics analysis, we were able to discover key resistance determinants, chromosomal mutations and mobile genetic elements that together contribute to the emergence of multidrug-resistant strains. The results show that resistance is not spontaneous but is a well-organized evolutionary phenomenon driven by a combination of selective pressure of antibiotics, acquisition of genes, and horizontal gene transfer.

The prevalence of β-lactamase, mecA genes, and vanA/vanB genes highlights the clinical obstacle of multidrug-resistant pathogens, and the close association between mobile genetic elements and resistance genes abundance validates the key role of plasmids, integrons, and transposons in enhancing resistance transmission. Phylogenetic analysis also showed lineage specific clustering of the resistant isolates, showing expansion of successful resistant strains clonally.

In general, this article highlights that antimicrobial resistance is an evolution-informed and genome-assisted process. Genomic surveillance in clinical microbiology is necessary to enhance diagnostic accuracy, monitor resistance development, and help to promote effective antimicrobial stewardship policies. These findings form a solid basis upon which special interventions can be created to counter the proliferation of resistance and preserve the effectiveness of subsequent antimicrobial treatment.

REFERENCES

[1] Baker, K. S., Jauneikaite, E., Nunn, J. G., Midega, J. T., Atun, R., Holt, K. E., ... & Peacock, S. J. (2023). Evidence review and recommendations for the implementation of genomics for antimicrobial resistance surveillance: reports from an international expert group. The Lancet Microbe, 4(12), e1035-e1039.

[2] Bandelt, H. J. (1995). Median-joining networks for inferring intraspecific phylogenies. Genetics, 141, 743.

[3] Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution, 16(1), 37-48.

[4] Bandelt, H. J., Macaulay, V., & Richards, M. (2000). Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Molecular phylogenetics and evolution, 16(1), 8-28.

[5] Consortium, H. M. P. (2012). Structure, function and diversity of the healthy human microbiome. Nature, 486(7402), 207-214.

[6] Deng, Q., Ramsköld, D., Reinius, B., & Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science, 343(6167), 193-196.

[7] Guest, J. F., Keating, T., Gould, D., & Wigglesworth, N. (2020). Modelling the annual NHS costs and outcomes attributable to healthcare-associated infections in England. BMJ open, 10(1), e033367.

[8] Pei, W., Feyerabend, T. B., Rössler, J., Wang, X., Postrach, D., Busch, K., ... & Rodewald, H. R. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature, 548(7668), 456-460.

[9] Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Developmental biology, 100(1), 64-119.

[10] Tan, Y., & Cahan, P. (2019). SingleCellNet: a computational tool to classify single cell RNA-Seq data across platforms and across species. Cell systems, 9(2), 207-213.

[11] Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., ... & Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature biotechnology, 32(4), 381-386.

[12] Trapnell, C. (2015). Defining cell types and states with single-cell genomics. Genome research, 25(10), 1491-1498.