Advances in Medical Sciences and Research
Vol. 1 (2026), No. 3, pp. 179-203
DOI: https://doi.org/10.55578/amsr.2607.012

REVIEW

Systems Biology, Multi-Omics, and Artificial Intelligence Paradigms in Bacterial Mutational Resistance and De Novo Antibacterial Drug Design

Monochura Saha [1,*], Miss Smriti [1], Priyanka Kumari [1]

[1] *Birla Institute of Technology, Mesra, Ranchi, Jharkhand 835215, India*

*Corresponding Author. Email Address: msaha@bitmesra.ac.in

Received: 2 June 2026, Revised: 29 June 2026, Accepted: 6 July 2026, Published: 13 July 2026

Abstract

Antimicrobial resistance (AMR) represents one of the most critical global threats to public health; it was associated with an estimated 4.95 million deaths in 2019 [1] and is projected to claim up to 10 million lives annually by 2050 if no effective interventions are implemented [2]. The rapid dissemination of multidrug-resistant (MDR) bacterial strains continues to render first-line and last-resort antibiotics ineffective, outpacing traditional drug discovery pipelines [3]. Historically, the development of new antibacterial agents relied on modifying existing chemical classes, a strategy that is increasingly vulnerable to rapid selection of resistance under clinical pressure [4]. To overcome this bottleneck, the modern research paradigm is shifting from single-molecule investigations to systems-level analyses [5]. The convergence of high-throughput multi-omics profiling spanning genomics, transcriptomics, proteomics, and metabolomics with advanced artificial intelligence (AI) and machine learning (ML) architectures offers a powerful framework for deciphering the complex molecular underpinnings of resistance and accelerating target-directed drug discovery [6].

Keywords: Antimicrobial Resistance (AMR), Multidrug Resistant (MDR), Artificial Intelligence (AI), Machine Learning (ML), Drug Design.

1. INTRODUCTION

In current healthcare science, antimicrobial resistance (AMR) is one of the most critical and rising global threats to public health, causing 1.27 million direct deaths in 2019 and associated with up to 4.95 million total deaths globally. According to landmark public health projections, this rising AMR threat could escalate drastically, claiming up to 10 million lives annually by 2050 if effective, system-level interventions are not integrated into medical science [7-8]. This rising AMR threat could claim up to 10 million lives annually by 2050 if no effective intervention is made in medical science [10,11]. The rapid global spread of multidrug-resistant (MDR) bacterial strains continues to render first-line and last-resort antibiotics ineffective, outpacing traditional drug discovery pipelines [12-14]. In clinical and public health settings, the AMR crisis is a continuous race against evolutionary times, where the selection pressure exerted by antibiotic usage outstrips the rate at which new therapies reach patients [15-18].

Historically, the development of new antibacterial agents relied on experimental screening or the systematic modification of existing chemical scaffolds, a simple modification strategy that is increasingly susceptible to rapid cross-resistance selection under clinical pressure [19-21]. This reductionist, single-molecule paradigm often treats drug targets and resistance mechanisms as isolated, static entities, overlooking the systemic complexity of bacterial physiology [22,23]. Conventional pipelines fail to capture

dynamic physiological adjustments, phenotypic tolerance, or the complex regulatory networks that bacteria deploy to survive. Consequently, this one-drug, one-target approach has led to high attrition rates, high costs, and a clinical pipeline that is dangerously thin relative to evolving resistance [23-25].

To overcome these compounding bottlenecks, the modern research paradigm is shifting from localized, single-molecule investigations to holistic, system-level analyses [26,27]. Biological entities do not function in isolation; rather, bacterial survival is governed by complex, interconnected networks of genes, transcripts, proteins, and metabolites [28-31]. Systems biology offers a quantitative framework for interrogating the coordinated interplay among genetic variants, gene expression profiles, protein-protein interactions, and downstream metabolic fluxes. By examining the bacterial cell as a unified dynamic system, researchers can identify critical metabolic chokepoints and regulatory hubs that are robust to mutational escape, paving the way for the design of evolution-resilient therapeutics [6,32,33].

This systemic transition is fundamentally powered by the revolution in high-throughput multi-omics profiling. While single-omics approaches, such as genomics, provide a valuable but static catalog of the bacterial resistome, they fail to reflect real-time adaptive states, transcriptional plasticity, or metabolic rewiring [34-36]. Integrating multiple omics layers, including genomics, transcriptomics, proteomics, and metabolomics, allows computational biologists to map the entire flow of biological information. This multi-layered perspective reveals how a single genomic mutation propagates through regulatory cascades to remodel the outer envelope or re-route central metabolic networks under drug pressure [35,37-39].

However, the sheer scale, high dimensionality, and extreme heterogeneity of multi-omics datasets present formidable computational barriers, exceeding the analytical capacity of classical statistical methods. Herein lies the transformative role of Artificial Intelligence (AI) and Machine Learning (ML). Modern AI/ML architectures ranging from classical interpretable classifiers to deep neural networks, graph neural networks, and transformer-based foundation models excel at extracting high-level abstractions and latent biological features from raw, noisy, and high-dimensional inputs [40-45]. Fusing these machine-learning workflows with constraint-based systems biology modeling enables the automated generation of actionable hypotheses. This synergistic combination accelerates early-stage target discovery, predicts the impact of resistance-conferring mutations with unparalleled speed, and optimizes de novo antibacterial drug design to neutralize pan-resistant pathogens [19,46-52].

2. MECHANISMS AND COMPUTATIONAL METHODOLOGIES

2.1. Understanding of molecular mechanisms of bacterial mutational resistance

Bacteria adapt to antibiotic pressure through two primary evolutionary strategies: the acquisition of foreign resistance determinants via horizontal gene transfer (HGT) and de novo mutations in native chromosomal genes [53]. While HGT mechanisms such as transformation, transduction, and conjugation efficiently disseminate resistance genes across diverse microbial communities, mutational resistance alters endogenous cellular structures, thereby compromising drug efficacy [53]. Spontaneous mutations occur naturally during DNA replication, but their selection and enrichment are greatly accelerated by sublethal antibiotic exposure, often due to premature cessation of treatment, suboptimal dosing, or environmental contamination [53,54]. Sub-inhibitory concentrations of bactericidal drugs can trigger stress responses that upregulate error-prone DNA polymerases, elevating the baseline mutation rate and fostering a genetic breeding ground for high-level resistance [53]. Mutational variations typically decrease bacterial susceptibility through three principal mechanisms [53].

a. **Modifications of Drug Targets:** Single-nucleotide polymorphisms (SNPs) or small insertions and deletions within target-encoding genes can induce conformational changes in the drug-binding pocket [53]. These alterations decrease the antibiotic's binding affinity while preserving the essential physiological function of the target enzyme (Table 1) [55].

Table 1. Enzymatic Inactivation for mutational adaptation

Resistance mechanism	Affected Gene(s) / Protein	Representative Pathogen(s)	Impacted Antibiotic Class(es)	Functional Outcome	Ref.
Target Modification	(b-subunit of RNA polymerase)	*Mycobacterium tuberculosis*, *Escherichia coli*	Rifamycins (Rifampin)	Decreased drug binding affinity and transcription capacity	[7]
Target Modification	mec-A (Penicillin-binding protein 2a)	*Staphylococcus aureus*	b-Lactams (Methicillin)	Low-affinity transpeptidase activity enabling cell wall synthesis	[53,55]
Porin Downregulation	opr-D	*Pseudomonas aeruginosa*	Carbapenems (Imipenem, Meropenem)	Reduced outer membrane permeability and decreased drug influx	[56]
Porin Modification	omp-K 36	*Klebsiella pneumoniae*	Cephalosporins, Carbapenems	Structural narrowing of the passive diffusion pore	[56]
Active efflux pump	mex B(RND Transporter)	*Pseudomonas aeruginosa*	Fluoroquinones, β-Lactams	Proton motive force-driven active extrusion of drugs	[56]
Active Efflux Pump	nor A (MFS Transporter)	*Staphylococcus aureus*	Fluoroquinolones	Increased drug extrusion across the cytoplasmic membrane	[56]
Enzymatic Inactivation	blaCTX-M-15,	Enterobacteriaceae	Cephalosporins, Carbapenems	Hydrolysis of the cyclic β-lactam ring structure	[56]
Enzymatic Modification	aac, aph, ant	*Staphylococcus aureus*, *Pseudomonas aeruginosa*	Aminoglycosides	Covalent modification (acetylation, phosphorylation)	[56]

b. **Reduction of Intracellular Drug Concentrations:** To reach their intracellular targets, hydrophilic antibiotics must cross the bacterial outer membrane, which Gram-negative pathogens restrict by

modifying outer membrane proteins [56]. Mutations that downregulate, structurally alter, or cause the complete loss of porins substantially decrease membrane permeability (Table-1) [56].

c. **Enzymatic Inactivation:** Mutational adaptation also accelerates the functional diversification of existing enzymes, enabling them to hydrolyze or chemically modify antibiotic structures [56]. Furthermore, aminoglycoside resistance is driven by aminoglycoside-modifying enzymes (AMEs), including acetyltransferases (AAC), phosphotransferases (APH) [53], and nucleotidyltransferases (ANT), which alter drug molecules to prevent target binding (Table 1) [56].

2.2. Multi-Omics Integration and Systems Biology Methodologies

Single-omics approaches, such as genomics, present a static view of the bacterial resistome and often fail to capture dynamic physiological adjustments, phenotypic tolerance, biofilm-mediated protection, or metabolic rewiring [3]. Systems biology addresses these gaps by integrating multiple molecular layers into a cohesive framework and tracking the flow of cellular information from DNA mutations to altered metabolic phenotypes.

This integrative approach combines: **(a) Genomics/Metagenomics:** Catalogs genetic variants, SNPs, and acquired resistance determinants; **(b) Transcriptomics:** Quantifies real-time gene expression shifts during drug exposure; **(c) Proteomics:** Evaluates protein abundance, post-translational modifications, and cellular translation capacity; **(d) Metabolomics/Exometabolomics:** Measures intracellular metabolites and extracellular foot printing to reflect active physiological states; **(e) Epigenomics:** Assesses DNA methylation patterns that govern phenotypic heterogeneity and mutation rates.

2.3. The three methodological strata of computational modeling

To process these vast, heterogeneous, and sparse multi-omics datasets, researchers organize AMR modeling into three methodological strata (Table 2a).

Table 2a. Comparison of different modeling

Classical Machine Learning	Structural & Deep Genomic Architectures	Transformer-Based & LLM Systems
Models: Random Forest, Support Vector Machines, Logistic Regression	Models: Convolutional Neural Networks, DeepARG, Graph Neural Networks	Models: Foundation genomic models, Large Language Models (LLMs)
Attributes: High interpretability, requires manual feature engineering	Attributes: Automated representation learning, network topology-aware	Attributes: Multimodal fusion, context-aware sequence learning

To guide clinicians and translational researchers in choosing optimal technical frameworks, especially within resource-constrained or time-sensitive settings, it is essential to analyze the structural trade-offs of these three computational strata. Table 2b evaluates their operational capabilities under clinical realities such as data scarcity, interpretability thresholds, and real-time computational needs.

Table 2b. Clinical Trade-offs and Selection Guidance for Computational AMR Strata

Technical Metric	Classical Machine Learning (RF, SVM, Logistic Regression)	Deep Learning Architectures (CNNs, RNNs, GNNs, DeepARG)	Transformer & Foundation Models (Geneformer, DNABERT, Nucleotide Transformer)	Ref.
Training Data Requirements	Performs well with small- to medium-	Requires substantially larger	Pretrained on billions of genomic or transcriptomic	[52, 58-65]

	sized datasets using engineered genomic or phenotypic features.	labeled datasets to optimize high-dimensional representations and reduce overfitting.	tokens; downstream fine-tuning can be performed using relatively small task-specific datasets through transfer learning.	
(**Performance Under Data Scarcity**	Generally robust when informative handcrafted features are available but may struggle to generalize beyond the training distribution.	Performance decreases under limited data because of overfitting unless regularization or augmentation strategies are employed.	Transfer learning and few-shot adaptation substantially improve performance in low-data regimes compared with training from scratch.	[64]
Interpretability	High. Feature importance, regression coefficients, and decision boundaries provide relatively transparent biological interpretation.	Moderate. Internal representations are difficult to interpret and generally require explainable AI methods such as SHAP, Integrated Gradients, or Grad-CAM.	Lower intrinsic interpretability. Attention mechanisms provide sequence-level associations but do not necessarily indicate causal biological relationships.	[43, 58, 62, 66, 67]
Computational Requirements	Low computational cost; training and inference can typically be performed on standard desktop computers.	Moderate computational demand; GPU acceleration is often beneficial for efficient training and inference.	Very high computational cost during pretraining, often requiring large GPU or TPU clusters; inference after pretraining is substantially less demanding.	[52, 58-66]
Ability to Capture Complex Biological Relationships	Limited ability to model higher-order nonlinear interactions unless explicitly engineered.	Effectively captures nonlinear sequence motifs, structural features, and complex genotype–phenotype relationships.	Learns long-range contextual dependencies and transferable biological representations across diverse genomic datasets.	[51-57-64]
Primary Limitations	Performance depends strongly on feature engineering and prior biological knowledge.	Large data requirements, increased computational complexity, and reduced interpretability.	High computational cost, limited interpretability, potential domain shift during fine-tuning, and ongoing challenges in biological validation.	[57-61,66]
Recommended Clinical Applications	Routine antimicrobial susceptibility prediction, hospital surveillance, and implementation in	Comprehensive genomic resistance prediction, resistance gene discovery, and genotype–phenotype	Population-scale genomic surveillance, multimodal clinical foundation models, transfer learning, and precision medicine applications.	[57-64]

	resource-limited diagnostic laboratories.	mapping in research laboratories.		

2.4. Classical and Interpretable Machine Learning

Algorithms such as Random Forests (RF), Support Vector Machines (SVM), and Logistic Regression (LR) utilize handcrafted genomic features (e.g., k-mer frequencies, gene presence-absence matrices, and known SNPs) to predict minimum inhibitory concentrations (MICs) and resistance classes. These models are computationally efficient and offer high interpretability through feature-importance scores, which help prioritize novel candidates for laboratory validation [67].

2.5. Structural and deep genomic architecture

Deep learning (DL) models bypass manual feature engineering by learning hierarchical, abstract representations directly from raw multi-omics sequences or structural inputs [20]. Neural architectures, such as DeepARG, use deep learning frameworks to categorize sequence reads into resistance classes with high precision [67]. Graph Neural Networks (GNNs), such as Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs), leverage biological network topology (e.g., protein-protein interaction networks from the STRING database) to model regulatory relationships, track network-level propagation, and identify key resistance hubs [68].

2.6. Transformer-Based and Foundation Artificial Intelligence Systems

Large language models (LLMs) and transformer architecture utilize self-attention mechanisms to learn long-range dependencies within protein sequences and genomic regions. These systems integrate genomic data with clinical metadata and epidemiological records to support automated surveillance, evolutionary forecasting, and real-time clinical decision support.

3. INTEGRATIVE MULTI-OMICS AND DEEP LEARNING FRAMEWORKS

3.1. Multi-Omics Integration Strategies and Deep Learning Architectures

A primary challenge in systems biology is integrating diverse biological data modalities, each with varying scales, dimensionality, and noise profiles. Artificial intelligence frameworks address these differences using three main data-fusion paradigms (Table 3) [69].

Table 3. Multi-Omics Integration Strategies

Integration Strategy	Definition & Mechanism	Key Advantages	Core Limitations	Representative ML/DL Architectures
Early Integration	Direct concatenation of raw features from different omics layers into a single high-dimensional input vector.	Simple to implement; retains all raw biological information.	Susceptible to the "curse of dimensionality"; struggles with disparate data distributions and noise.	Concatenated Neural Networks, Classical Classifiers (RF, SVM, LR) [70]
Late Integration	Training independent models on individual	Preserves the unique topological structures and	Ignores non-linear, cross-layer synergistic regulatory	Ensemble Classifiers, Voting Classifiers, Deep

	omics layers and aggregating their predictions (e.g., via voting or averaging).	properties of each independent omics layer.	interactions and feedback loops.	Learning Predictors [69]
Intermediate Integration	Fuses low-dimensional, latent representations extracted from separate neural network branches within a joint model.	Captures non-linear crosstalk across layers; resolves complex biological pathways.	Highly complex; requires matched multi-omics cohorts and substantial computational resources.	Multi-Branch Deep Architectures, Graph Neural Networks (GNNs), Autoencoders [69]

3.2. Synergizing Genome-Scale Metabolic Models with Machine Learning

Genome-scale metabolic models (GEMs) represent structured, curated knowledge bases that translate annotated genetic sequences into biochemically consistent stoichiometric matrices (s-matrices) [71]. Within this matrix, columns represent enzymatic reactions and rows represent metabolite concentrations, linked by the steady-state mass balance assumption:

$$S \cdot v = 0 \quad \text{(i)}$$

where ט is the vector of metabolic fluxes [71]. While Flux Balance Analysis (FBA) optimizes cellular objective functions (e.g., growth rate via biomass accumulation functions), classic constraint-based methods struggle with biological degeneracy, in which multiple distinct flux distributions yield the same objective value [71].

By using machine learning and high-throughput multi-omics data to copy GEMs, researchers can refine these solutions to deliver condition-specific phenotypic predictions.

3.3. Mechanistic Constraints and Solution Space Optimization

Multi-omics integration refines GEM feasibility regions by applying distinct constraint logics across biological layers.

a. **Transcriptomes as Context-Specific Switches:** Rather than mapping linearly to flux magnitudes, transcript levels act as context filters [72]. Algorithms such as GIMME, iMAT, and MADE use gene expression thresholds to prune inactive metabolic pathways and define active network topologies under specific environments (Figure 1) [72].

b. **Proteomes as capacity valves and budgets:** Proteomic data dictates absolute enzymatic capacity limits v_{max} [72]. By integrating protein abundance (E) with enzyme turnover numbers (k_{cat}), enzyme-constrained models scale maximum reaction velocities using the thermodynamic limit:

$$v_i \leq k_{\text{cat},i} \cdot E_i$$

This approach accounts for cellular proteome allocation trade-offs and energetic costs [72].

c. **Metabolomes as Thermodynamic Anchors:** Intracellular metabolomics data impose physical boundaries on flux directions by defining Gibbs free energies (DG) of reactions, ensuring that flux occurs only along thermodynamically favorable pathways [72].

d. **Fluxomics as Calibration Targets:** Experimental 13C-flux measurements provide standard values to validate model predictions, reducing the degeneracy of FBA solution spaces [72].

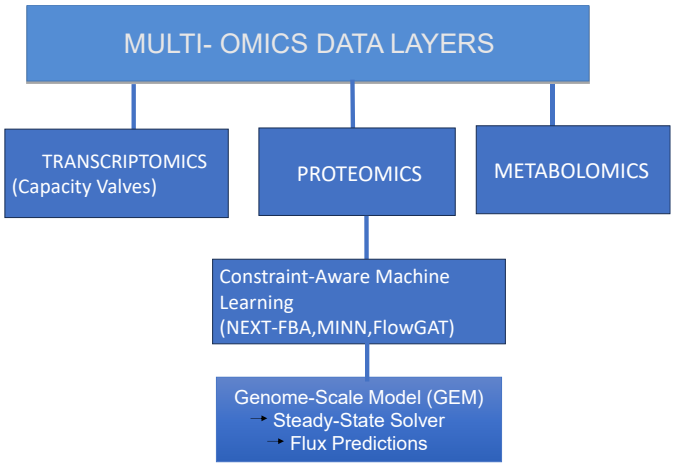

Figure 1. Schematic illustration of multiomics data analysis

4. COMPUTATIONAL SURROGATE MODELS AND KINETIC PREDICTIONS

4.1. Neural Surrogates and Kinetic Parameter Predictions

A key bottleneck in constraint-based systems biology is the scarcity of experimentally validated enzyme kinetic parameters (k_{cat}, KM, and KI), particularly for non-model pathogens [72]. Deep learning tools (such as DLKcat, TurNuP, UniKP, CatPred, and RealKcat) address this gap by predicting key kinetic constants directly from protein sequences, 3D structural models, and substrate chemical descriptors [72]. These ML-derived kinetic values are integrated as confidence-weighted priors into enzyme-constrained metabolic models [72].

Furthermore, computational biologists deploy constraint-aware neural surrogates, such as NEXT-FBA and Metabolic Integration Neural Networks (MINN) [72]. These architectures embed the biochemical stoichiometric network topology directly into neural network regularization layers [72]. By enforcing physical conservation laws during training, these models achieve excellent out-of-distribution (OOD) generalization, allowing researchers to predict complex flux distributions, gene essentiality, and target susceptibility under antibiotic stress (Figure 1) [72].

5. PRACTICAL IMPLEMENTATIONS: COMPREHENSIVE CASE STUDIES

To demonstrate how integrating systems biology, multi-omics, and artificial intelligence accelerates AMR mutation detection and drug discovery, this section analyzes nine recent case studies (Figure 2).

5.1. Multiscale Pangenomic Machine Learning across ESKAPE Pathogens

Researchers developed a multiscale machine learning framework to analyze pangenomes across approximately 13,000 bacterial isolates sourced from the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) (Figure 2) [67].

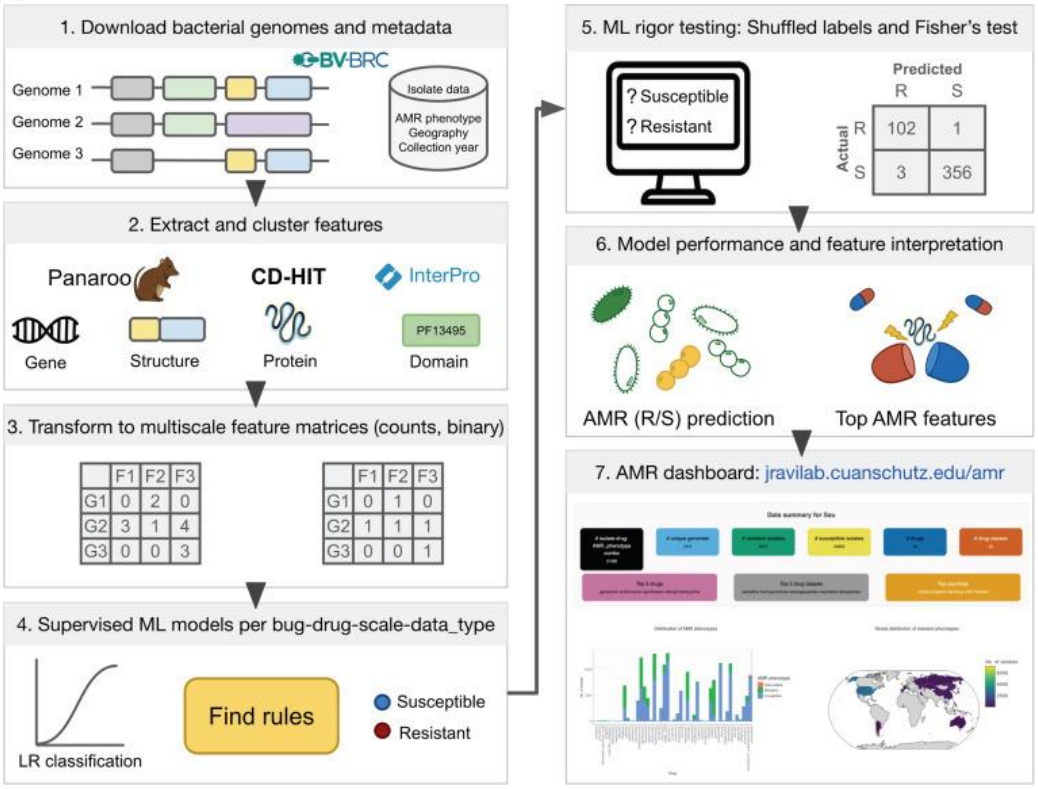

Figure 2: The author described a bacterial species-agnostic workflow that leverages publicly available datasets and supervised machine learning (ML) approaches to predict antimicrobial resistance (AMR) phenotypes of bacterial isolates against specific antibiotics. Briefly, the workflow begins with: 1) retrieval of genomic (.fna) and annotated protein (.faa) sequences, along with corresponding AMR phenotype labels and metadata, from the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) for selected bacterial species; 2) clustering of features using Panaroo (for genes and gene structures) and CD-HIT (for proteins), followed by extraction of protein domains using InterPro; 3) generation of per-genome count-based and presence/absence matrices for ML analysis; 4) training of supervised logistic regression (LR) models for each species/drug/feature scale/data type combination to classify drug resistance or susceptibility; 5) establishment of baseline controls by shuffling labels within each matrix and retraining ML models to remove biological signals, alongside non-ML analysis using Fisher's exact test on molecular features; 6) comparison of model performance across bacterial species, antibiotics, molecular scales, and data types, with further analysis of the most predictive molecular features associated with AMR; and 7) dissemination of the results through an interactive data dashboard (jravilab.org/amr) to facilitate exploration, visualization, and hypothesis generation[67].

a. **Pathogens Investigated:** The ESKAPE panel (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Enterobacter* species) [67].
b. **Computational Strategy:** Genomic sequencing data were processed to construct species-specific pangenomesv[68]. Rather than relying on known resistance databases, the framework extracted pangenomic features across multiple biological scales, including gene presence-absence matrices, protein sequences, structural domains, and gene neighborhood rearrangements. These features were mapped to species-agnostic Clusters of Orthologous Groups (COGs) to capture broader evolutionary signatures. Linear classifiers (Elastic Net and Logistic Regression) were trained to predict phenotypic resistance [67].
c. **Performance Metrics:** The models achieved a median normalized Matthews correlation coefficient (nMCC) of 0.89. Crucially, the classifiers maintained their performance when evaluated on independent

geographical and temporal holdout sets, demonstrating resilience to distribution shifts.

d. **Key Findings:** The framework successfully identified canonical resistance markers, such as the *mecA* gene in methicillin-resistant *S. aureus* (MRSA) and *the ermC gene* in macrolide-resistant isolates. Additionally, it highlighted novel non-canonical features such as hypothetical proteins and structural genomic motifs within mobile genetic elements, providing candidates for target discovery [67].

5.2. DIABLO Multi-Omics Modeling of Staphylococcus aureus Serum Adaptation [28]

To survive within the bloodstream, *S. aureus* must coordinate its central carbon metabolism, iron transport, and oxidative stress defenses to overcome host nutritional immunity. The integrated model revealed that *S. aureus* survival in human serum relies on metabolic transitions to gluconeogenesis and the TCA cycle, utilizing alternative carbohydrates (such as mannose and fructose) and host amino acids (specifically glutamate, proline, arginine, and histidine). This metabolic shift was linked to the activation of virulence factors, including gamma-hemolysin genes (*hlgA*, *hlgB*, *hlgC*) and immune evasion molecules (*ecb*, *sbi*, *vraX*). Phenotypic validation of isogenic mutants confirmed that the DIABLO-identified genes *gapdhB*, *sucA*, *sirA*, *sstD*, and *perR* are critical for *S. aureus* survival in human serum (Figure 3) [28].

a. **Pathogens Investigated:** Five distinct clinical isolates of *S. aureus* (BPH2760, BPH2819, BPH2900, BPH2947, and BPH2986).[28].

b. **Computational Strategy:** Researchers performed integrated transcriptomic, proteomic, and metabolomic profiling of clinical strains exposed to human serum versus glucose-rich control medium (RPMI-1640) [28]. Rather than relying on late integration, the investigators utilized **DIABLO** (*Data Integration Analysis for Biomarker discovery using Latent Components*), a supervised multiblock integration method based on sparse partial least squares discriminant analysis (sPLS-DA). DIABLO was optimized to select a minimal set of multi-omic features that co-vary across data types and distinguish exposure conditions (Figure 3).

c. **Performance Metrics:** Extracted multi-omics signatures showed strong correlation across layers, demonstrating a highly coordinated biological response to human serum. The final DIABLO model identified 60 highly discriminative features [28].

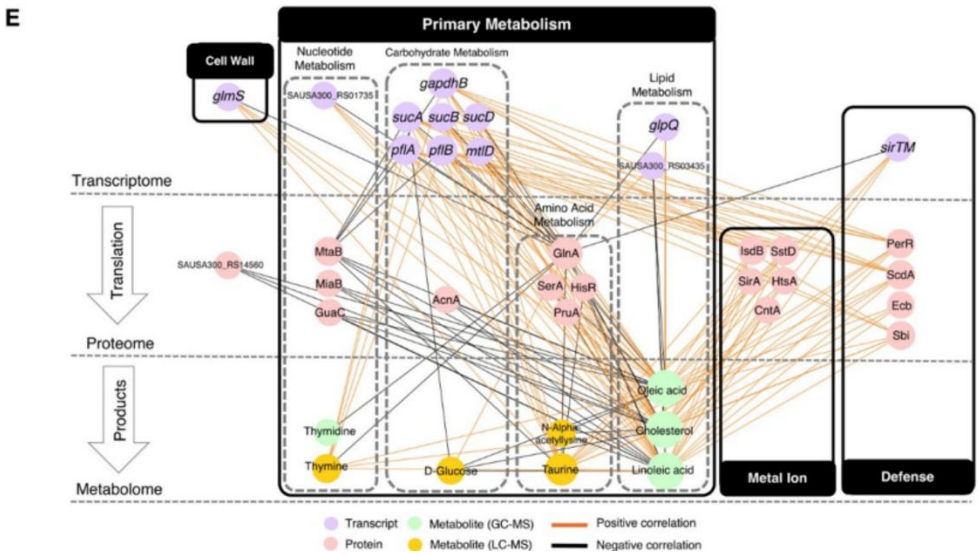

Figure 3: Multi-omics identification of molecular signatures associated with *S. aureus* adaptation to serum. (A) Schematic representation of the multi-omics workflow employed to identify signature molecules of *S. aureus* in response to human serum exposure. (B) Unsupervised analysis performed using principal component analysis (PCA). The PCA sample plot demonstrates consistent clustering patterns across the different omics datasets, including transcriptomics, proteomics, and metabolomics (GC-MS and LC-MS), with samples grouped according to culture conditions (SERUM and RPMI). (C) Supervised analysis conducted using multiblock sparse partial least squares discriminant analysis (sPLS-DA). The diagnostic sample scatter plot illustrates separation by medium (SERUM and RPMI). Correlation coefficients between the first components of each dataset are shown as numerical values below the diagonal in each block. (D) Loading plot showing the molecular features identified by PLS-DA from the first component of each dataset as the most discriminatory between experimental conditions. (E) Correlation network displaying pairwise associations (> |0.9|) among transcripts, proteins, and metabolites. Network edges represent correlations, with orange and black edges denoting positive and negative correlations, respectively. Node size corresponds to the number of interactions, while node colors indicate dataset type: purple for transcripts, pink for proteins, green for GC-MS metabolites, and yellow for LC-MS metabolites [28].

5.3. Transcriptome-Based Predictive Modeling in Pseudomonas aeruginosa

Pseudomonas aeruginosa features extensive genomic plasticity, enabling rapid mutational resistance under therapeutic pressure.

a. **Pathogens Investigated:** 410 multidrug-resistant clinical isolates of *Pseudomonas aeruginosa* [73].
b. **Computational Strategy:** Researchers profiled the transcriptomes of clinical isolates and trained ten different machine learning classifiers on gene expression (GEXP) features to predict susceptibility to ceftazidime, ciprofloxacin, meropenem, and tobramycin [73].
c. **Performance Metrics**: The Random Forest and gradient-boosted ensembles outperformed other models, with the tobramycin susceptibility model achieving a predictive accuracy of 98.8% and a sensitivity of 100% [73].
d. **Key Findings:** The model identified specific transcriptomic signatures linked to resistance, highlighting alterations in biofilm formation, membrane transport, amino acid metabolism, and virulence. Crucially, the model identified a core signature of ten mRNAs that were consistently deregulated across all four antibiotic classes, pointing to a shared transcriptional program governing multidrug resistance. These shared transcripts are associated with oxidative stress defense, iron acquisition, and the activation of active efflux pumps (such as the *mexB* system) [73].

5.4. De Novo Broad-Spectrum Antibiotic Discovery via Deep Learning

Traditional high-throughput chemical screening is costly and slow, and it tends to identify derivatives

of known drug classes. To address this, deep learning was deployed to screen massive chemical spaces for structurally novel compounds (Figure 4).[74].

a. **Pathogens Investigated:** *Acinetobacter baumannii, Mycobacterium tuberculosis*, and *Clostridioides difficile*.
b. **Computational Strategy:** Researchers trained a deep neural network on the chemical structures and experimental growth inhibition profiles of a large, diverse compound library. Deep models learned complex structure-activity relationships beyond classical quantitative structure-activity relationships (QSAR) descriptors, focusing on features that inhibit growth while remaining structurally distinct from those of clinical antibiotics. The model was then used to screen *in silico* drug libraries containing over 100 million molecules.
c. **Performance Metrics:** The model prioritized candidate compounds, leading to the identification of Halicin. Halicin demonstrated potent bactericidal activity against multiple pan-resistant clinical pathogens, successfully clearing infections with minimal toxicity in murine models (Figure 4) [74].

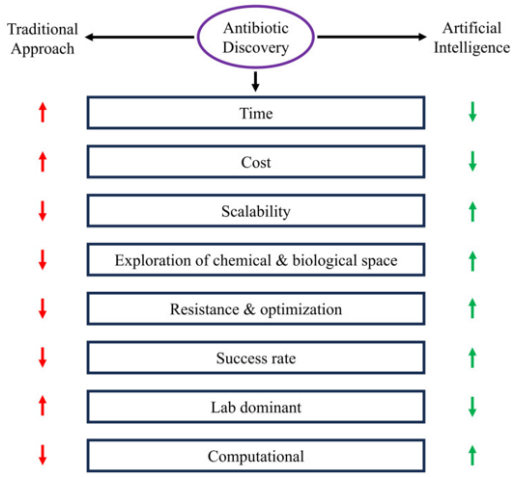

Figure 4: Relative comparison of conventional antibiotic discovery approaches and AI-driven methodologies. The figure highlights the key parameters influencing antibiotic development, including time, cost, scalability, exploration of chemical and biological space, resistance management and optimization, success rate, and dependence on laboratory-based versus computational workflows. Conventional discovery strategies are generally associated with longer development timelines, higher costs, limited scalability, and substantial reliance on wet-lab experimentation. In contrast, AI-based approaches enable faster, more cost-effective discovery processes, greater scalability, enhanced exploration of chemical and biological space, superior resistance-aware optimization, and higher overall success rates through computationally guided decision-making [74].

d. **Key Findings:** Mechanistic evaluation revealed that Halicin functions through an alternative mode of action: it dissipates the transmembrane proton motive force, disrupting electrochemical gradients across bacterial inner membranes and preventing ATP synthesis [74].

5.5. Deep Recurrent Architectures for Anti-Tuberculosis Antimicrobial Peptide Design

Antimicrobial peptides (AMPs) represent promising alternatives to traditional small-molecule therapies due to their rapid membrane-disruptive activity and low potential to select for resistance mutations. However, designing pathogen-specific AMPs is constrained by the scarcity of target-specific training datasets.

a. **Pathogens Investigated:** *Mycobacterium tuberculosis (Mtb)*.
b. **Computational Strategy:** Researchers developed a deep learning protocol that coupled Long Short-Term Memory (LSTM) networks with transfer learning. The model was pre-trained on broad AMP databases to capture general features of antimicrobial function, and subsequently fine-tuned on a limited, highly curated dataset of Mtb-active peptides. In parallel, an LSTM-based generative model was constructed to design de novo, TB-specific peptide sequences.

c. **Performance Metrics:** Among the evaluated variants, a unidirectional LSTM with a frozen encoder achieved the highest classification performance on a held-out TB test set, yielding an accuracy of 90% and an Area Under the ROC Curve (AUROC) of 0.97. The generative algorithm produced 100 de novo sequences, 94 of which were predicted as highly active by independent AMP classifiers (Tablen4) [75].

5.6. Hybrid Quorum Sensing and Machine Learning Synthetic Biology Systems

The clinical utility of conventional antibiotics is often undermined by rapid mutational resistance and off-target toxicity. Synthetic biology platforms address these issues by engineering microbial circuits that coordinate metabolic behavior, but these systems require dynamic adjustment under changing environments.

a. **Pathogens Investigated:** Engineered therapeutic strains targeting opportunistic pathogens.
b. **Computational Strategy**: Researchers constructed hybrid systems combining Quorum Sensing (QS) genetic circuits with machine learning algorithms. The QS circuits monitored cell density and signaling molecules, while the machine learning models utilized multi-omics data platforms to predict metabolic outputs and optimize circuit parameters in real time.
c. **Performance Metrics:** The hybrid QS-ML systems achieved real-time, autonomous feedback control over gene expression and metabolic outputs, maintaining circuit stability across varying environments.
d. **Key Findings:** The integration of ML-enabled engineered systems to dynamically optimize therapeutic peptide secretion and metabolic outputs, thereby reducing selective pressure for resistance mutations and improving targeting precision against pathogenic populations (Table-4) [76].

5.7. MOMLIN and Multimodal Biomarker Extraction Frameworks

Identifying robust, multi-modal biomarkers that predict therapeutic response represents a major challenge due to high dimensionality and sparse datasets.

a. **Pathogens/Framework Focus:** Pathogen microenvironments and drug-response classification.
b. **Computational Strategy:** Researchers developed MOMLIN (*Multi-modal and -omics Machine Learning Integration*), a computational framework that utilizes weighted multi-class sparse canonical correlation analysis (WMSCCA). MOMLIN employs an adaptive weighting scheme across pairwise SCCA models, preventing high-dimensional layers (such as genomics or transcriptomics) from dominating low-dimensional layers (such as metabolomics) during feature extraction. This allows the model to isolate class-specific sparse components and multimodal biomarkers.
c. **Performance Metrics:** MOMLIN achieved a classification performance with an average AUROC of 0.989, outperforming conventional integration methods (including standard Sparse Canonical Correlation Analysis and Multi-Omics Factor Analysis) by at least 10% [76].

6. COMPARATIVE PERFORMANCE OF SYSTEMS BIOLOGY AND MACHINE LEARNING CASE STUDIES

To compare the diverse methodologies, data modalities, and performance metrics across these implementations, the following table summarizes the case studies discussed (Table 4).

Table 4. Comparative Performance of Systems Biology and Machine Learning Case Studies

Case Study	Target Pathogen / System	Integrated Omics Layers	Computational / AI Framework	Validation Metrics & Key	Ref.
1. Multiscale Pangenomic ML	ESKAPE Pathogens	Genomics, structural protein domains, pangenome neighborhoods	Elastic Net, Logistic Regression, COG mapping	Median n = MCC. Successfully isolated canonical and novel AMR	[77]

				markers; resilient to spatial/temporal holdouts.	
2. DIABLO Serum Adaptation	*Staphylococcus aureus*	Transcriptomics, Proteomics, Metabolomics (GC-MS/LC-MS)	DIABLO (supervised sparse multiblock PLS-DA)	Strong cross-layer correlation (r>1). Confirmed roles of *gapdhB*, *sucA*, *sirA*, *sstD*, and *perR* via mutants.	[28]
3. Transcriptome-Based GEXP	*Pseudomonas aeruginosa*	Transcriptomics	Random Forest, Gradient-Boosted Trees, SVM	Tobramycin model achieved 98.8% accuracy and 100% sensitivity. Isolated 10 core multidrug resistance mRNAs.	[73]
4. De Novo Halicin Screening	*A. baumannii*, *M. tuberculosis*, *C. difficile*	Deep molecular representations, high-throughput phenotypic screen	Deep Neural Network for virtual screening	Screened >100 M molecules. Discovered **Halicin**, which acts by dissipating transmembrane proton motive force.	[74]
5. Recurrent AMP Generation	*Mycobacterium tuberculosis*	Peptide sequence chemistry, validated AMP datasets	LSTM recurrent network with Transfer Learning	Frozen encoder model achieved 90% accuracy (AUC =). Generative candidates shared ≥ identity with clinical AMPs.	[75]
6. Exometabolomic Footprinting	*Pseudomonas aeruginosa*	Untargeted Extracellular Metabolomics,	Dynamic mathematical modeling of	Dynamic pathways tracked 1,921	[78]

		Time-Kill Kinetics	exometabolites	features. L-arginine and L-ornithine correlated with killing (r = 0.82).	
7. Hybrid QS-ML Systems	Synthetic Therapeutic Circuits	Synthetic gene circuits, Multi-omics data platforms	Adaptive Machine Learning coupled with genetic circuits	Real-time, autonomous feedback control of gene expression, reducing selective pressure for mutational escape.	[76]
8. MOMLIN Framework	Drug Response Classification	Genomics, Transcriptomics, Epigenomics, Proteomics	Weighted Multi-Class SCCA (WMSCCA)	Achieved an average AUROC =, outperforming standard SCCA and MOFA integration methods by ≥.	[79]
9. MIA Tensor Framework	Clinical Patient Phenotypes	Genomics, Transcriptomics, Proteomics, Epigenomics	Tensor Decomposition, Fuzzy C-Means, Supervised Random Forest	Outperformed SNF and PINS. Identified clinical subtypes and isolated 167 biomarker genes.	[80]

6.1. Benchmarking Datasets, Reproducibility, and Open-Source Infrastructure

A primary barrier to moving multi-omics machine learning from standalone academic publications to verified clinical pipelines is cross-study reproducibility. Heterogeneous pipeline inputs, varying filtering parameters, and distinct data formats often degrade performance when models transition between diagnostic environments. To build trust and cross-validation integrity, the community relies on standardized benchmarking repositories and unified workflow frameworks. Table 5 outlines the core open-access infrastructure that drives current peer-reviewed computational AMR pipelines.

Table 5. Standardized Benchmarking Databases, Repositories, and Reproducibility Infrastructure

Category	Resource / Platform	Primary Curated	Role in Computational AMR Research	Ref.

		Content		
Benchmarking Databases	**BV-BRC (Bacterial & Viral Bioinformatics Resource Center)**	Whole-genome assemblies, isolate metadata, antimicrobial susceptibility testing (AST) data, MIC values, and pathogen annotations	Supports standardized training and external validation of genome-based AMR prediction models across clinically relevant bacterial species.	[81]
	CARD (Comprehensive Antibiotic Resistance Database)	Expert-curated antibiotic resistance genes (ARGs), resistance mechanisms, mutations, ontologies, and the Resistance Gene Identifier (RGI)	Provides a high-quality reference database for AMR gene annotation, resistance mechanism classification, and benchmarking of computational prediction algorithms.	[82]
	NCBI AMRFinderPlus	Curated resistance genes, virulence factors, stress response genes, and resistance-associated point mutations	Widely used standardized reference tool for genomic AMR detection and comparison with machine-learning and deep-learning prediction pipelines.	[83]
	ResFinder	Acquired antimicrobial resistance genes and chromosomal resistance mutations	Enables standardized identification of resistance determinants from bacterial genome sequences and serves as a commonly used baseline in comparative studies.	[84]
Open-Source Computational Frameworks	**Chemprop**	Graph neural network (message-passing neural network) framework for molecular property prediction	Facilitates AI-driven discovery and prediction of antimicrobial compound activity against resistant pathogens using chemical structure representations.	[85]
	mixOmics (DIABLO)	Statistical and machine-learning framework for	Enables integrative biomarker discovery and feature selection across heterogeneous biological datasets relevant to host–pathogen	[86]

		multi-omics integration (transcriptomics, proteomics, metabolomics, microbiome data)	interactions and AMR mechanisms.	
Reproducible Workflow Systems	**Docker / Singularity (Apptainer)**	Containerized computational environments including software, dependencies, and operating system libraries	Ensures computational reproducibility, software portability, and consistent execution of bioinformatics pipelines across computing platforms.	[88, 89]
	Nextflow	Workflow management system supporting scalable, container-native computational pipelines	Enables reproducible, parallel, and portable execution of complex multi-omics and genomic AMR workflows across local, HPC, and cloud infrastructures.	[89]
	Snakemake	Python-based workflow management framework with provenance tracking	Facilitates reproducible data processing, scalable workflow automation, and transparent pipeline documentation for computational biology research.	[90]

7. OPERATIONAL CHALLENGES, TRANSLATIONAL BARRIERS, AND FUTURE OUTLOOK

The clinical and pharmaceutical translation of hybrid systems biology and artificial intelligence architectures is significantly constrained by the complexity of *in vivo* microenvironments. Rather than progressing instantly from exposure to genetic resistance, pathogens under selective drug pressure undergo several distinct, pre-genetic physiological stages that can confound standard predictive models.

Stage 1: Metabolic Compensation: This initial stage occurs within 5 to 30 minutes of antibiotic exposure. Bacteria initiate stress response pathways and rewire central carbon pathways to maintain redox and energetic homeostasis. This phase is characterized by rapid fluctuations in energy indices (ATP/ADP and NADH/NAD+ ratios) and transient rerouting of carbon flow. It is completely reversible upon drug washout, and durable envelope remodeling does not occur at this timescale [91].

Stage 2: Sub-lethal Adaptation: This intermediate phase occurs over 30 to 240 minutes of continued sub-inhibitory exposure. Pathogens remodel their cellular membrane composition, such as by enriching cardiolipin or phosphatidylglycerol, and upregulate active efflux pumps. While these changes remain pre-genetic and largely reversible, they slow down killing kinetics, resulting in phenotypic tolerance or persistence [91].

Stage 3: Genetic Resistance: Sustained therapeutic stress selects for de novo mutations or horizontal gene transfer events. These genetic changes are permanent and structurally fixed, altering the minimum inhibitory concentration (MIC) and establishing high-level clinical resistance (Figure 5).

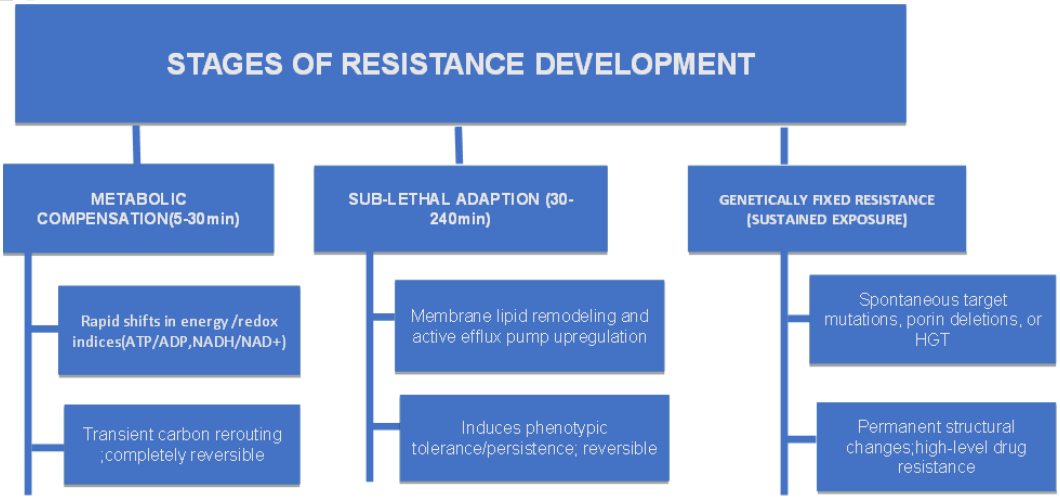

Figure 5. Schematic description of antibacterial resistance development

7.1. Operational Challenges and Clinical Constraints

To achieve routine clinical applicability, systems-level computational architectures must overcome several analytical bottlenecks

a. **Data Sparsity and Technical Noise:** High-throughput multi-omics profiling is often limited by small matched-sample cohort sizes, technical artifacts, and platform variability. Batch effects and missing data points can compromise model performance, requiring robust imputation algorithms and standardized preprocessing workflows.

b. **The Interpretability Bottleneck:** Deep learning architectures, such as multi-branch neural networks and transformers, function as complex black boxes. This lack of interpretability can limit their utility for identifying novel therapeutic targets or understanding causal molecular mechanisms. Integrating Explainable AI (XAI) with constraint-based metabolic models is essential for providing biologically meaningful predictions [72].

c. **Distribution Shifts and Robustness:** AI models trained on localized datasets often suffer from performance degradation when evaluated on new geographic or temporal populations. Establishing robust, multicenter validation protocols and deploying federated learning models are critical to ensure generalizability.

7.2. A Stepwise Roadmap for Bench-to-Bedside Clinical Translation

Moving AI predictions and multi-omics biomarkers from digital discovery loops to bedside diagnostic routines requires a systematic, multi-phase clinical translation framework. Figure 6 details the five essential developmental milestones required to safely and effectively implement computational AMR tools at the point of care.

Stage 1: *In silico* Refinement and Pre-Clinical Safety Profiling: Raw computational hits (such as de novo small molecules or antimicrobial peptides generated by generative models) must immediately pass through secondary algorithmic filters. These steps predict absorption, distribution, metabolism, excretion, and toxicity (ADMET) parameters. Deep neural classifiers score potential compound candidates against human cellular toxicity models, hemolytic indices, and structural alerts for mutagenicity or cardiotoxicity. This stage narrows down millions of virtual library compounds into a manageable subset of highly viable, synthesizable chemical structures [92-94].

Stage 2: Standardized Preclinical Wet-Lab Validation Cascades: Candidate therapeutics or diagnostic multi-omic biomarkers move from the computer into high-throughput wet-lab screening. This phase evaluates candidate compounds using automated microdilution screening assays against panels of WHO priority pathogens. To better replicate real-world clinical infections, researchers run these tests using advanced *in vitro* multi-lineage organoid models or microfluidic organs-on-a-chip. This step captures multi-

layer bacterial responses (such as metabolic compensation or transient phenotypic tolerance) within physiological microenvironments. Top candidates are then validated *in vivo* using mouse thigh or lung infection models to verify pharmacokinetic and pharmacodynamic (PK/PD) safety and efficacy [19,95].

Stage 3: Regulatory Compliance and Standardization Gateways: To secure regulatory clearance (such as FDA 510(k) approval or CE marking), the computational model must be designed in accordance with the regulatory guidelines governing Software as a Medical Device (SaMD). This phase requires locking down the model's weights and architecture, eliminating potential data leakage, and demonstrating clear performance stability against clinical distribution shifts. Furthermore, multi-omics data preparation pipelines must secure international laboratory certifications (such as CLIA and CAP). This ensures that automated extraction protocols, sequencing depths, and quality control steps are perfectly uniform across different clinical pathology sites [96].

Stage 4: Adaptive, Stratified Clinical Trials: Traditional clinical trial designs often struggle to account for the wide genetic diversity observed in real-world bacterial infections. Next-generation translational frameworks utilize adaptive trial strategies. Here, real-time genomic sequencing and machine learning stratification models group patients based on the specific resistance profile of their infecting pathogen. Instead of prescribing a broad, one-size-fits-all antibiotic regimen, trials use machine learning predictions to assign targeted, narrow-spectrum therapies. This agile structure accelerates clinical endpoint confirmation while significantly protecting patients from ineffective, toxic drug exposures [97].

Stage 5: Bedside Integration and Clinical Decision Support (CDS) Systems: The final phase of translation automates these tools directly at the clinical point of care. Computational architectures are deployed as automated microservices integrated within institutional Electronic Health Records (EHR) via secure interoperability protocols (such as HL7 FHIR). When a hospital laboratory uploads a pathogen's quick sequencing profile, an integrated cloud microservice calculates phenotypic resistance probabilities within minutes. The tool then delivers a clear report directly to the attending physician’s screen, suggesting targeted, safe, and highly personalized antibiotic choices. This complete workflow empowers real-time, evidence-based antimicrobial stewardship right at the bedside. [98]

8. CONCLUSION

The integration of systems biology with comprehensive biological datasets represents a fundamental shift in how we understand and combat antibiotic resistance. By moving away from isolated, single-molecule targets and embracing the entire flow of genetic, protein, and metabolic information, researchers can now map how complex cellular networks adapt to drug-induced stress. This system-level approach, powered by advanced predictive software and constraint-based metabolic models, successfully bridges the gap between digital predictions and physical laboratory testing. Ultimately, translating these computational workflows into clinical settings requires standardized validation protocols, collaborative data sharing, and unified public health surveillance systems. By aligning predictive technology with real-world biology, this modern research paradigm offers a scalable, sustainable pathway to bypass emerging resistance and discover next-generation, life-saving therapeutics.

Ethics Approval and Consent to Participate: Not applicable.

Consent for publication: Not applicable.

Competing Interests: The author declares no competing interests.

Funding: No external funding was received for this work.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author Contributions: MCS conceived the idea, reviewed the literature, and wrote and edited the manuscript. MS and PK edited the figures, table, and the manuscript.

Acknowledgments: The author expresses gratitude to the Birla Institute of Technology, Mesra, for providing a platform for conducting and publishing impactful research and review articles. During the preparation of this manuscript, the author(s) used ChatGPT and Google Gemini (OpenAI) for language refinement and summarizing references. The authors reviewed and edited the output as needed and took full responsibility for the content of the published article.

REFERENCES

[1] Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., … Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet*, *399*(10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

[2] Naghavi, M., Vollset, S. E., Ikuta, K. S., Swetschinski, L. R., Gray, A. P., Wool, E. E., … Murray, C. J. (2024). Global burden of bacterial antimicrobial resistance 1990–2021: A systematic analysis with forecasts to 2050. *The Lancet*, *404*(10459), 1199–1226. https://doi.org/10.1016/S0140-6736(24)01867-1

[3] Hossain, E., & Yousefi, N. (2026). Computational paradigms for antimicrobial resistance prediction: Integrating multi-omics, structural modeling, and foundation artificial intelligence systems. *Briefings in Bioinformatics*, *27*(3), bbag219. https://doi.org/10.1093/bib/bbag219

[4] Yow, H.-Y., Govindaraju, K., Lim, A. H., & Abdul Rahim, N. (2022). Optimizing antimicrobial therapy by integrating multi-omics with pharmacokinetic/pharmacodynamic models and precision dosing. *Frontiers in Pharmacology*, *13*, 915355. https://doi.org/10.3389/fphar.2022.915355

[5] Shen, Y., Zhang, P., Luo, J., Chen, S., Gu, S., Lin, Z., … Yang, L. (2026). Artificial intelligence drives advances in multi-omics analysis and precision medicine for sepsis. *Biomedicines*, *14*(2), 261. https://doi.org/10.3390/biomedicines14020261

[6] Francine, P. (2022). Systems biology: New insight into antibiotic resistance. *Microorganisms*, *10*(12), 2362. https://doi.org/10.3390/microorganisms10122362

[7] O'Neill, J. (2014). *Antimicrobial resistance: Tackling a crisis for the health and wealth of nations*. Review on Antimicrobial Resistance.

[8] Fernandez, G. (2022). Turning the juggernaut [Editorial]. *The Lancet Infectious Diseases*, *22*(5), E75–E76. https://doi.org/10.1016/S1473-3099(22)00254-3

[9] de Kraker, M. E., Stewardson, A. J., & Harbarth, S. (2016). Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS Medicine*, *13*(11), e1002184. https://doi.org/10.1371/journal.pmed.1002184

[10] Pulingam, T., Parumasivam, T., Gazzali, A. M., Sulaiman, A. M., Chee, J. Y., Lakshmanan, M., … Sudesh, K. (2022). Antimicrobial resistance: Prevalence, economic burden, mechanisms of resistance and strategies to overcome. *European Journal of Pharmaceutical Sciences*, *170*, 106103. https://doi.org/10.1016/j.ejps.2021.106103

[11] Lewis, K. (2020). The science of antibiotic discovery. *Cell*, *181*(1), 29–45. https://doi.org/10.1016/j.cell.2020.02.056

[12] Ribeiro da Cunha, B., Fonseca, L. P., & Calado, C. R. (2019). Antibiotic discovery: Where have we come from, where do we go? *Antibiotics*, *8*(2), 45. https://doi.org/10.3390/antibiotics8020045

[13] Lewis, K. (2013). Platforms for antibiotic discovery. *Nature Reviews Drug Discovery*, *12*(5), 371–387. https://doi.org/10.1038/nrd3975

[14] Michael, C. A., Dominey-Howes, D., & Labbate, M. (2014). The antimicrobial resistance crisis: Causes, consequences, and management. *Frontiers in Public Health*, *2*, 145. https://doi.org/10.3389/fpubh.2014.00145

[15] Salam, M. A., Al-Amin, M. Y., Salam, M. T., Pawar, J. S., Akhter, N., Rabaan, A. A., & Alqumber, M. A. (2023). Antimicrobial resistance: A growing serious threat for global public

health. *Healthcare*, *11*(13), 1946. https://doi.org/10.3390/healthcare11131946
[16] Makabenta, J. M. V., Nabawy, A., Li, C.-H., Schmidt-Malan, S., Patel, R., & Rotello, V. M. (2021). Nanomaterial-based therapeutics for antibiotic-resistant bacterial infections. *Nature Reviews Microbiology*, *19*(1), 23–36. https://doi.org/10.1038/s41579-020-0420-3
[17] Hetta, H. F., Ramadan, Y. N., Al-Harbi, A. I., Ahmed, E. A., Battah, B., Abd Ellah, N. H., … El-Saber Batiha, G. (2023). Nanotechnology as a promising approach to combat multidrug resistant bacteria: A comprehensive review and future perspectives. *Biomedicines*, *11*(2), 413. https://doi.org/10.3390/biomedicines11020413
[18] Miranda, R. R., Parthasarathy, A., & Hudson, A. O. (2022). Exploration of chemical biology approaches to facilitate the discovery and development of novel antibiotics. *Frontiers in Tropical Diseases*, *3*, 845469. https://doi.org/10.3389/fitd.2022.845469
[19] Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., … Collins, J. J. (2020). A deep learning approach to antibiotic discovery. *Cell*, *180*(4), 688–702. https://doi.org/10.1016/j.cell.2020.01.021
[20] Sakenova, N., Cacace, E., Orakov, A., Huber, F., Varik, V., Kritikos, G., … Typas, A. (2025). Systematic mapping of antibiotic cross-resistance and collateral sensitivity with chemical genetics. *Nature Microbiology*, *10*(1), 202–216.
[21] Maura, D., Ballok, A. E., & Rahme, L. G. (2016). Considerations and caveats in anti-virulence drug development. *Current Opinion in Microbiology*, *33*, 41–46. https://doi.org/10.1016/j.mib.2016.07.015
[22] Zampieri, M., Zimmermann, M., Claassen, M., & Sauer, U. (2017). Nontargeted metabolomics reveals the multilevel response to antibiotic perturbations. *Cell Reports*, *19*(6), 1214–1228. https://doi.org/10.1016/j.celrep.2017.04.028
[23] Prasad, N. K., Seiple, I. B., Cirz, R. T., & Rosenberg, O. S. (2022). Leaks in the pipeline: A failure analysis of gram-negative antibiotic development from 2010 to 2020. *Antimicrobial Agents and Chemotherapy*, *66*(5), e00054-22. https://doi.org/10.1128/aac.00054-22
[24] Bakkeren, E., Diard, M., & Hardt, W.-D. (2020). Evolutionary causes and consequences of bacterial antibiotic persistence. *Nature Reviews Microbiology*, *18*(9), 479–490. https://doi.org/10.1038/s41579-020-0361-x
[25] Wei, Y., Palacios Araya, D., & Palmer, K. L. (2024). Enterococcus faecium: Evolution, adaptation, pathogenesis and emerging therapeutics. *Nature Reviews Microbiology*, *22*(11), 705–721. https://doi.org/10.1038/s41579-024-01056-2
[26] Farha, M. A., French, S., & Brown, E. D. (2021). Systems-level chemical biology to accelerate antibiotic drug discovery. *Accounts of Chemical Research*, *54*(8), 1909–1920. https://doi.org/10.1021/acs.accounts.0c00771
[27] Roemer, T., & Boone, C. (2013). Systems-level antimicrobial drug and drug synergy discovery. *Nature Chemical Biology*, *9*(4), 222–231. https://doi.org/10.1038/nchembio.1210
[28] Mujchariyakul, W., Walsh, C. J., Giulieri, S., Cramond, C., Lê Cao, K.-A., Stinear, T. P., … Monk, I. R. (2026). Integrated multi-omics reveals coordinated Staphylococcus aureus metabolic, iron transport, and stress responses to human serum. *mSystems*, *11*(3), e01183-25. https://doi.org/10.1128/msystems.01183-25
[29] Cloots, L., & Marchal, K. (2011). Network-based functional modeling of genomics, transcriptomics and metabolism in bacteria. *Current Opinion in Microbiology*, *14*(5), 599–607. https://doi.org/10.1016/j.mib.2011.07.022
[30] Typas, A., & Sourjik, V. (2015). Bacterial protein networks: Properties and functions. *Nature Reviews Microbiology*, *13*(9), 559–572. https://doi.org/10.1038/nrmicro3483
[31] Klosik, D. F., Grimbs, A., Bornholdt, S., & Hütt, M.-T. (2017). The interdependent network of gene regulation and metabolism is robust where it needs to be. *Nature Communications*, *8*(1), 534. https://doi.org/10.1038/s41467-017-00493-5
[32] Sarker, M., Talcott, C., & Galande, A. K. (2013). *In silico* systems biology approaches for the identification of antimicrobial targets. In *In silico Models for Drug Discovery* (pp. 13–30). Humana

Press. https://doi.org/10.1007/978-1-62703-342-8_2

[33] Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biology*, *20*(1), 121. https://doi.org/10.1186/s13059-019-1730-3

[34] Porreca, V., Barbagallo, C., Corbella, E., Peres, M., Stella, M., Mignogna, G., … Ferro, A. (2024). Unveil intrahepatic cholangiocarcinoma heterogeneity through the lens of omics and multi-omics approaches. *Cancers*, *16*(16), 2889. https://doi.org/10.3390/cancers16162889

[35] Sakaguchi, T., Irifune, Y., Kamada, R., & Sakaguchi, K. (2025). Bacterial systematic genetics and integrated multi-omics: Beyond static genomics toward predictive models. *International Journal of Molecular Sciences*, *26*(19), 9326. https://doi.org/10.3390/ijms26199326

[36] Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. Ø. (2009). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, *7*(2), 129–143. https://doi.org/10.1038/nrmicro2089

[37] Young, A. T., Carette, X., Helmel, M., Steen, H., Husson, R. N., Quackenbush, J., … Mattila, J. T. (2021). Multi-omic regulatory networks capture downstream effects of kinase inhibition in Mycobacterium tuberculosis. *NPJ Systems Biology and Applications*, *7*(1), 8. https://doi.org/10.1038/s41540-020-00155-9

[38] Acierno, C., Barletta, F., Nevola, R., Rinaldi, L., Sasso, F. C., Adinolfi, L. E., … Marrone, A. (2025). Metabolic rewiring of bacterial pathogens in response to antibiotic pressure—a molecular perspective. *International Journal of Molecular Sciences*, *26*(12), 5574. https://doi.org/10.3390/ijms26125574

[39] Vemuri, G. N., & Aristidou, A. A. (2005). Metabolic engineering in the-omics era: Elucidating and modulating regulatory networks. *Microbiology and Molecular Biology Reviews*, *69*(2), 197–216. https://doi.org/10.1128/MMBR.69.2.197-216.2005

[40] Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics data integration, interpretation, and its application. *Bioinformatics and Biology Insights*, *14*, 1177932219899051. https://doi.org/10.1177/1177932219899051

[41] Misra, B. B., Langefeld, C., Olivier, M., & Cox, L. A. (2019). Integrated omics: Tools, advances and future approaches. *Journal of Molecular Endocrinology*, *62*(1), R21–R45. https://doi.org/10.1530/JME-18-0122

[42] Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. *Briefings in Bioinformatics*, *18*(5), 851–869. https://doi.org/10.1093/bib/bbw068

[43] Ma, Y., Yao, A., Chen, X., Wang, L., Ma, C., Xi, X., … Wang, X. (2021). Generation of truncated derivatives through *in silico* enzymatic digest of peptide GV30 target MRSA both *in vitro* and *in vivo*. *Computational and Structural Biotechnology Journal*, *19*, 4984–4996. https://doi.org/10.1016/j.csbj.2021.08.032

[44] Gaudelet, T., Day, B., Jamasb, A. R., Soman, J., Regep, C., Liu, G., … Taylor, S. (2021). Utilizing graph machine learning within drug discovery and development. *Briefings in Bioinformatics*, *22*(6), bbab159. https://doi.org/10.1093/bib/bbab159

[45] Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., … Zitnick, C. L. (2021). Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, *118*(15), e2016239118. https://doi.org/10.1073/pnas.2016239118

[46] Vaidyanathan, R., Kalishwaralal, K., Gopalram, S., & Gurunathan, S. (2009). Nanosilver—The burgeoning therapeutic molecule and its green synthesis [Retracted article]. *Biotechnology Advances*, *27*(6), 924–937. https://doi.org/10.1016/j.biotechadv.2009.06.002

[47] Chavali, A. K., D'Auria, K. M., Hewlett, E. L., Pearson, R. D., & Papin, J. A. (2012). A metabolic network approach for the identification and prioritization of antimicrobial drug targets. *Trends in Microbiology*, *20*(3), 113–123. https://doi.org/10.1016/j.tim.2011.12.002

[48] Yuan, Y., & Bar-Joseph, Z. (2019). Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Sciences*, *116*(52), 27151–

27158. https://doi.org/10.1073/pnas.1914536116
[49] Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, *18*(1), 83. https://doi.org/10.1186/s13059-017-1215-1
[50] Nembrini, S., König, I. R., & Wright, M. N. (2018). The revival of the Gini importance? *Bioinformatics*, *34*(21), 3711–3718. https://doi.org/10.1093/bioinformatics/bty373
[51] Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., & Zhang, L. (2018). DeepARG: A deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*, *6*(1), 23. https://doi.org/10.1186/s40168-018-0401-z
[52] Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. *Nature*, *624*(7990), 80–85. https://doi.org/10.1038/s41586-023-06735-9
[53] Munita, J. M., & Arias, C. A. (2016). Mechanisms of antibiotic resistance. In *Virulence Mechanisms of Bacterial Pathogens* (pp. 481–511). American Society of Microbiology. https://doi.org/10.1128/9781555819286.ch17
[54] Misra, R. (2026). Mutation-based mechanisms of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, e00101-25. https://doi.org/10.1128/mmbr.00101-25
[55] Al-Wahaibi, L. H., Alzahrani, H. A., Bräse, S., Youssif, B. G., & El-Sheref, E. M. (2026). Design and synthesis of novel quinolin-2-ones as dual inhibitors of DNA gyrase and topoisomerase IV possessing antibacterial and apoptotic antiproliferative effects. *Journal of Molecular Structure*, 146525. https://doi.org/10.1016/j.molstruc.2026.146525
[56] Nusrat, S., Aliyu, M., & Zohora, F. T. (2025). Mechanisms of antimicrobial resistance: From genetic evolution to clinical manifestations. *AIMS Microbiology*, *11*(4), 1007.
[57] Su, M., Satola, S. W., & Read, T. D. (2019). Genome-based prediction of bacterial antibiotic resistance. *Journal of Clinical Microbiology*, *57*(3), e01405-18. https://doi.org/10.1128/JCM.01405-18
[58] Nguyen, M., Long, S. W., McDermott, P. F., Olsen, R. J., Olson, R., Stevens, R. L., … Musser, J. M. (2019). Using machine learning to predict antimicrobial MICs and associated genomic features for nontyphoidal Salmonella. *Journal of Clinical Microbiology*, *57*(2), e01260-18. https://doi.org/10.1128/JCM.01260-18
[59] Yang, Y., Niehaus, K. E., Walker, T. M., Iqbal, Z., Walker, A. S., Wilson, D. J., … Clifton, D. A. (2018). Machine learning for classifying tuberculosis drug-resistance from DNA sequencing data. *Bioinformatics*, *34*(10), 1666–1671. https://doi.org/10.1093/bioinformatics/btx708
[60] Van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The third revolution in sequencing technology. *Trends in Genetics*, *34*(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.006
[61] Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Al Sayed, Z. R., Hill, M. C., … Ellinor, P. T. (2023). Transfer learning enables predictions in network biology. *Nature*, *618*(7965), 616–624. https://doi.org/10.1038/s41586-023-06239-6
[62] Ji, Y., Zhou, Z., Liu, H., & Davuluri, R. V. (2021). DNABERT: Pre-trained Bidirectional Encoder Representations from Transformers model for DNA-language in genome. *Bioinformatics*, *37*(15), 2112–2120. https://doi.org/10.1093/bioinformatics/btab083
[63] Dalla-Torre, H., Gonzalez, L., Mendoza-Revilla, J., Lopez Carranza, N., Grzywaczewski, A. H., Oteri, F., … Bencherif, T. (2025). Nucleotide transformer: Building and evaluating robust foundation models for human genomics. *Nature Methods*, *22*(2), 287–297.
[64] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., … Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint*, arXiv:2108.07258. https://arxiv.org/abs/2108.07258
[65] Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K.-R. (2019). *Explainable AI: Interpreting, explaining and visualizing deep learning*. Springer Nature. https://doi.org/10.1007/978-3-030-28954-6
[66] Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks.

In *Proceedings of the 34th International Conference on Machine Learning* (pp. 3319–3328). PMLR.
[67] Ghosh, A., Brenner, E. P., Vang, C. K., Wolfe, E. P., Boyer, E. A., Lesiyon, R. L., … Allen, H. K. (2025). From sequence to signature: Machine learning uncovers multiscale feature landscapes that predict AMR across ESKAPE pathogens. *bioRxiv*. https://doi.org/10.1101/2025.07.03.663053
[68] Hasan, M. I., Smyth, D. S., & Teufel, A. I. (2025). Network analysis of antimicrobial resistance in Staphylococcus aureus: Characterization of hub genes and their functional implications. *NAR Genomics and Bioinformatics*, *7*(4), lqaf147. https://doi.org/10.1093/nargab/lqaf147
[69] Ren, Y., Bai, H., Wang, J., Yang, Y., & Wang, Y. (2026). Deep learning-enabled multi-omics integration: A new frontier in precise drug target discovery. *Biology*, *15*(5), 410. https://doi.org/10.3390/biology15050410
[70] Vora, N., Shah, S., Patel, P., & Shah, M. (2025). Artificial intelligence and multi-omics in drug discovery: A deep learning-powered revolution. *Cure & Care*, 100011.
[71] Fang, X., Lloyd, C. J., & Palsson, B. O. (2020). Reconstructing organisms *in silico*: Genome-scale models and their emerging applications. *Nature Reviews Microbiology*, *18*(12), 731–743. https://doi.org/10.1038/s41579-020-00463-9
[72] Shahreen, N., Osinuga, A., Malla, S., Razmpour, T., Tabibian, M., & Saha, R. (2026). Multi-omics integration in genome-scale metabolic models: A review of constraint-based approaches. *Molecular Omics*, *22*(2), aaiag005.
[73] Kula, C., Erguven, I., Ozcelik, B., Gulfidan, G., & Arga, K. Y. (2026). Transcriptome-based machine learning models to predict antimicrobial resistance in Pseudomonas aeruginosa. *OMICS: A Journal of Integrative Biology*, *30*(2), 94–104.
[74] Sanapalli, B. K. R., Palit, S., Deshpande, A., Tokala, R., Sigalapalli, D. K., & Sanapalli, V. (2026). Artificial intelligence and the discovery of antibiotics: Reinventing with opportunities, challenges, and clinical translation. *Antibiotics*, *15*(2), 233. https://doi.org/10.3390/antibiotics15020233
[75] Wang, L., Campino, S., Clark, T. G., & Phelan, J. E. (2025). Long short-term memory-based deep learning model for the discovery of antimicrobial peptides targeting Mycobacterium tuberculosis. *Bioinformatics Advances*, *5*(1), vbaf274. https://doi.org/10.1093/bioadv/vbaf274
[76] Tuan, D. A., Uyen, P. V. N., Ho, T., & Vo Thien Nhan, N. (2025). Hybrid quorum sensing and machine learning systems for precision gene regulation: Revolutionizing synthetic biology and autonomous therapeutic applications. *Frontiers in Microbiology*, *16*, 1520227.
[77] Alquaddoomi, F. S., Burke, J. T., Sosinski, L., Mayer, D. A., Brenner, E. P., Chen, S. Z., … Allen, H. K. (2022). MolEvolvR: A web-app for characterizing proteins using molecular evolution and phylogeny. *bioRxiv*. https://doi.org/10.1101/2022.02.18.461833
[78] Tait, J. R., Anderson, D., Nation, R. L., Creek, D. J., & Landersdorfer, C. B. (2024). Identifying and mathematically modeling the time-course of extracellular metabolic markers associated with resistance to ceftolozane/tazobactam in Pseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy*, *68*(4), e01081-23. https://doi.org/10.1128/aac.01081-23
[79] Rashid, M. M., & Selvarajoo, K. (2024). Advancing drug-response prediction using multi-modal and-omics machine learning integration (MOMLIN): A case study on breast cancer clinical data. *Briefings in Bioinformatics*, *25*(4), bbae300. https://doi.org/10.1093/bib/bbae300
[80] Zhang, Y., Liu, L., Liu, Z., Liu, Q., Ma, L., & Zhang, Z. (2025). Machine learning for integrative multi-omics clustering and feature gene identification. *bioRxiv*. https://doi.org/10.1101/2025.11.28.691099
[81] Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., … von Mering, C. (2023). The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, *51*(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000
[82] Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., … McArthur, A. G. (2023). CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, *51*(D1), D690–D699. https://doi.org/10.1093/nar/gkac989

[83] Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J. G., Haendiges, J., Haft, D. H., … Klimke, W. (2021). AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Scientific Reports*, *11*(1), 12728. https://doi.org/10.1038/s41598-021-91456-0

[84] Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., … Aarestrup, F. M. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *Journal of Antimicrobial Chemotherapy*, *75*(12), 3491–3500. https://doi.org/10.1093/jac/dkaa345

[85] Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., … Barzilay, R. (2019). Analyzing learned molecular representations for property prediction. *Journal of Chemical Information and Modeling*, *59*(8), 3370–3388. https://doi.org/10.1021/acs.jcim.9b00237

[86] Rohart, F., Gautier, B., Singh, A., & Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Computational Biology*, *13*(11), e1005752. https://doi.org/10.1371/journal.pcbi.1005752

[87] Merkel, D. (2014). Docker: Lightweight Linux containers for consistent development and deployment. *Linux Journal*, *239*(2), 2.

[88] Kurtzer, G. M., Sochat, V., & Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS ONE*, *12*(5), e0177459. https://doi.org/10.1371/journal.pone.0177459

[89] Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, *35*(4), 316–319. https://doi.org/10.1038/nbt.3820

[90] Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., van Dyken, P. C., Tomkins-Tinch, C. H., … Köster, J. (2025). Sustainable data analysis with Snakemake. *F1000Research*, *10*, 33. https://doi.org/10.12688/f1000research.29032.2

[91] Hussein, M., Creek, D. J., Baker, M., Rao, G. G., Li, J., & Velkov, T. (2025). Metabolomics-driven prediction of antibiotic resistance: A perspective on pre-genetic intervention. *npj Antimicrobials and Resistance*, *3*(1), 101.

[92] Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., … Zhang, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, *18*(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5

[93] Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022). Artificial intelligence in biological sciences. *Life*, *12*(9), 1430. https://doi.org/10.3390/life12091430

[94] Timmons, P. B., & Hewage, C. M. (2020). HAPPENN is a novel tool for hemolytic activity prediction for therapeutic peptides which employs neural networks. *Scientific Reports*, *10*(1), 10869. https://doi.org/10.1038/s41598-020-67945-5

[95] Theuretzbacher, U., Outterson, K., Engel, A., & Karlén, A. (2020). The global preclinical antibacterial pipeline. *Nature Reviews Microbiology*, *18*(5), 275–285. https://doi.org/10.1038/s41579-019-0288-6

[96] Mehta, V., Komanduri, A., Bhadouriya, R. S., Mehta, V., Johnson, M. D., Shrestha, P., … Reddy, S. (2025). Evaluating transparency in AI/ML model characteristics for FDA-reviewed medical devices. *npj Digital Medicine*, *8*(1), 673.

[97] Yang, H., & Zhao, L. (2025). *Clinical trial modernization: Technological, operational, and regulatory advances*. Chapman and Hall/CRC.

[98] Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine*, *3*(1), 17. https://doi.org/10.1038/s41746-019-0221-x