Poster Number: 1

Integrating Molecular Dynamics, Molecular Docking, and Machine Learning for Predicting SARS-CoV-2 Papain-like Protease Binders

Ann Varghese[1], Jie Liu[1], Tucker A Patterson[1], Huixiao Hong[1]

1. National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, USA

Background: Despite advances in vaccine development, the availability of effective antivirals for SARS CoV-2 remains limited. Among the druggable targets of SARS CoV-2, Papain-like protease (PLpro) is of particular interest due to its dual functions in viral polyprotein processing and host immune evasion. However, PLpro remains an underexplored target in the context of drug repurposing for COVID-19 treatment. To address this, we combined machine learning, molecular dynamics, and molecular docking to develop a predictive model that can distinguish PLpro binders from non-binders.

Methods: Ligand-bound SARS-CoV-2 PLpro structures were downloaded from the Protein Data Bank. The structures were then subjected to 1 µs molecular dynamics simulations, and the resulted trajectories were clustered into three groups using the K-means clustering algorithm in AMBER 16. A representative structure was selected for subsequent molecular docking. A Random Forest (RF) model was trained on docking results and evaluated using leave-one-out cross-validation.

Results: Docking of the FDA-approved drugs to PLpro showed a wide range of docking scores, indicating the varied binding potential of different drugs to PLpro. The RF model achieved >75% accuracy in identifying true PLpro binders and non-binders in the leave-one-out cross-validation.

Conclusions: Our findings demonstrate that integrating computational chemistry and machine learning techniques can provide an efficient and rapid framework for drug repurposing.

Keywords: SARS CoV-2 PLpro, Papain-like protease, Drug repurposing, Binding, Molecular Docking

Poster Number: 2

GanCtrl: A Generative AI Approach to Derive Study-Aligned Synthetic Controls for Reducing Concurrent Control Animal Use

Mansi Chandra[1,2], Xi Chen[1], Ting Li[1], and Weida Tong[1]

1. National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, USA
2. University of Arkansas at Little Rock and University of Arkansas for Medical Sciences Joint Bioinformatics Program, Little Rock, Arkansas 72204, USA

Background: Minimizing the use of animals as concurrent controls in *in vivo* studies directly supports the principles of the 3Rs (Replacement, Reduction, and Refinement). Current approaches primarily use historical data to construct virtual control groups (VCGs), but these methods are limited by cross-study variability arising from differences in study design, laboratory practices, and data heterogeneity.

Methods: We propose GanCtrl, a generative AI approach based on generative adversarial networks (GANs), to infer study-specific control conditions directly from time-matched treatment data. By generating synthetic controls intrinsically aligned with each study's biological, temporal, and technical context, GanCtrl avoids biases associated with cross-study normalization in traditional VCG methods. GanCtrl was applied to rat repeat-dose toxicity studies to simulate 38 clinical pathology endpoints under control conditions using corresponding treatment-group data.

Results: Synthetic controls closely approximated real concurrent controls. Differences between synthetic and real controls were smaller than both intra-laboratory and inter-laboratory baseline variations and were comparable to biological replicate variability. For practical utility, toxicity assessments using synthetic controls were compared with those using real concurrent controls, pooled intra-laboratory controls, and pooled inter-laboratory controls. Concordance between synthetic and real controls was approximately 72% for seven liver biomarkers and 68% for seven kidney biomarkers, comparable to or exceeding pooled control approaches. Importantly, synthetic controls preserved drug-induced clinical pathology changes (e.g., ALT elevation) and biologically relevant biomarker relationships (e.g., ALT–AST), consistent with established literature.

Conclusions: GanCtrl provides a generative AI-based methodology to generate study-aligned synthetic controls, offering a practical strategy to reduce reliance on concurrent control animals in toxicology studies.

Keywords: 3Rs, concurrent control groups (CCGs), virtual control groups (VCGs), generative AI, generative adversarial networks (GANs), toxicology

Poster Number: 3

Mobile Phones as Reservoirs for Bacterial Contamination: A Systematic Meta-Analysis

Nari Gu[1]

1. Department of Biology, University of Arkansas at Little Rock, Little Rock, AR, United States

Background: Mobile phones (MPs) are continuously handled in both healthcare and community settings, enabling frequent bacterial transfer between hands, skin, and device surfaces. Consequently, MPs have been recognized as potential fomites and reservoirs of clinically relevant bacteria. However, despite continuous research on MP contamination from 2005 to 2025, existing studies remain fragmented by region and methodology, and no synthesis has integrated contamination prevalence, bacterial diversity, and potential public health risks.

Methods: We conducted a systematic meta-analysis of 327 studies encompassing 39,987 MPs. Data were extracted to evaluate contamination rate, taxonomic composition at genus and species levels, and clinical relevance of identified isolates. Species co-occurrence networks were constructed, and pan-species accumulation curves were fitted using Heap's law to assess whether the MP-associated microbiome represents an open or closed community.

Results: Overall, 31,748 devices (79.4%) were contaminated, with significantly higher contamination in healthcare settings than in non-healthcare environments. Staphylococcus aureus was the most frequently isolated species, followed by opportunistic pathogens including Escherichia coli, Pseudomonas aeruginosa, Klebsiella pneumoniae, and Staphylococcus epidermidis. Of 240 species, many were potential human pathogens, and pathogenic taxa dominated MPs in both settings. Network analysis revealed a densely interconnected structure with hubs including S. aureus, E. coli, Acinetobacter baumannii, and P. aeruginosa. Pan-species modeling indicated an "open" community, suggesting that additional bacterial taxa may be detected with further sampling.

Conclusions: MPs function as widespread, pathogen-rich reservoirs and highlight the need for standardized device hygiene protocols, targeted behavioral interventions, and expanded bacterial surveillance to reduce MP-associated microbial transmission in both healthcare and public settings.

Keywords: Mobile phone, microbial contamination, cross-contamination, retrospective analysis, pathogen-disease network, pan-species

Poster Number: 4

Copper Resistance Drives Resistance to Naturally Occurring Aminoglycoside in Salmonella enterica isolated from food

Ashraf Khan[1], Ahmed, F. Hikal [1], Faisal A. Khan [1], Shaohua Zhao[1], Steven L. Foley[1]

1. Division of Microbiology, National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, USA

Background: The aim of this study was to examine the link between copper resistance and aminoglycoside resistance in Salmonella enterica (S. enterica).

Methods: We previously disrupted the acquired copper-resistance pco cassette in nontyphoidal Salmonella Senftenberg CVM N32755 through deletion of pcoABCD genes using allelic exchange mutagenesis. We also evaluated pco+ S. Anatum CVM N31410. Aminoglycosides susceptibility was assessed using both the microdilution Alamar Blue assay and Etest, with the minimum inhibitory concentrations (MICs) reported based on the microdilution Alamar Blue assay, except for the MICs of S. Senftenberg ΔpcoABCD and its parent strain against kanamycin, tobramycin, and gentamicin, which were determined using Etest. Bacterial membrane permeability was assayed using propidium iodide (PI) staining.

Results: In this report, we demonstrated that deletion of the copper-resistance genes (pcoABCD) in S. Senftenberg increased susceptibility to tobramycin, kanamycin, streptomycin, gentamicin, and spectinomycin. MIC assays showed a more than 10-fold increase in susceptibility to kanamycin, and spectinomycin, a structurally related aminocyclitol antibiotic, in the S. Senftenberg ΔpcoABCD mutant compared to the parent strain. Additionally, the mutant exhibited 4- and 20-fold increases in susceptibility to streptomycin and tobramycin, respectively. However, there was no increase in susceptibility to amikacin, netilmicin, or plazomicin. Remarkably, S. Anatum N31410 that carried the pco and sil copper-resistance clusters, but lacks known aminoglycoside-resistance genes, is resistant to tobramycin (MIC = 125 μg/mL), gentamicin (MIC = 620 μg/mL), kanamycin (MIC = 156 μg/mL), streptomycin (MIC ≥ 2000 μg/mL), and spectinomycin (MIC = 620 μg/mL).

Conclusions: These findings uncover a previously unrecognized linkage between copper-resistance and resistance to aminoglycosides, which suggests cross-resistance mediated by the pco cluster to naturally occurring aminoglycosides. These results provide novel insight into the crucial link between copper-resistance mechanism and antimicrobial resistance, which are of significant relevance to the scientific community.

Disclaimer: This article reflects the views of its authors and does not necessarily reflect those of the U.S. Food and Drug Administration. Any mention of commercial products is for clarification only and is not intended as approval, endorsement, or recommendation.

Keywords: Salmonella enterica, Copper resistance, Aminoglycoside

Poster Number: 5

Assessing Quality of Drug Nomenclatures in RxNorm to Enhance Drug Safety Surveillance using Unstructured Data

Wenjing Guo[1], Huixiao Hong[1]

1. National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, US

Adverse drug events are harms associated with drug use, whether the drug is used correctly or incorrectly. Identifying adverse drug events is vital in pharmacovigilance to safeguard public health. Drug safety surveillance can be performed using unstructured data. A comprehensive and accurate list of drug names is essential for effective identification of adverse drug events. While there are numerous sources for drug names, RxNorm is widely recognized as a leading resource. However, its effectiveness for unstructured data analysis in drug safety surveillance has not been thoroughly assessed. To address this, we evaluated the drug names in RxNorm for their suitability in unstructured data analysis and developed a refined set of drug names. Initially, we removed duplicates, the names exceeding 199 characters, and those that only describe administrative details. Drug names with four or fewer characters were analyzed using 18,000 drug-related PubMed abstracts to remove names which rarely appear in unstructured data. The remaining names, which ranged from five to 199 characters, were further refined to exclude those that could lead to inaccurate drug counts in unstructured data analysis. We compared the efficiency and accuracy of the refined set with the original RxNorm set by testing both on the 18,000 drug-related PubMed abstracts. The results showed a decrease in both computational cost and the number of false drug names identified. Further analysis of the removed names revealed that most originated from only one of the 14 sources. Our findings suggest that the refined set can enhance drug identification in unstructured data analysis, thereby improving pharmacovigilance.

Keywords: Adverse drug events, RxNorm, drug safety, drug name, surveillance, unstructured data

Poster Number: 6

Integrated Multi-Model Analysis Reveals Intestinal Barrier and Immune Alterations Induced by Anti-retroviral drug

Yaswanthi Yanamadala[1], Kuppan Gokulan[1], Chandra Mohan Reddy Muthumula[1], Vicki Sutherland[2], Helen Cunny[2], Janine H. Santos[2], Kelly Davis[3], Sangeeta Khare[1]

1. Division of Microbiology, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR
2. Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, NC
3. Toxicologic Pathology Associates, Jefferson, AR, United States

Background: Tri-combination antiretroviral therapy (TC-ART; abacavir/dolutegravir/lamivudine) is widely used during pregnancy to prevent vertical HIV transmission. However, its long-term effects on gut health remain unclear.

Methods: We integrated multi-model approaches (in vivo, in vitro, and NAM/organ-on-chip systems) with multi-omics analyses, including microbiome sequencing, targeted gene expression profiling (Qiagen panels), and multiplex cytokine analysis, to characterize intestinal epithelial, microbiome, and immune responses to TC-ART.

Results: In vivo, gestational and lactational exposure resulted in dose- and sex-dependent dysbiosis in aged rat offspring [1-3]. Mucosa-associated microbiota showed increased species richness in high-dose females, whereas males exhibited enrichment of opportunistic taxa. Despite minimal changes in fecal short-chain fatty acids (SCFAs), males demonstrated greater vulnerability, including reduced weight gain, villar atrophy, inflammatory dysregulation, and downregulation of permeability-related genes, while females showed compensatory gene responses with fewer cytokine alterations. In vitro, TC-ART induced concentration- and time-dependent epithelial dysfunction in T84 cells, with moderate exposure (500µM) causing partial barrier loss and higher concentrations (≥2000µM) leading to rapid disruption. Junctional gene downregulation co-occurred with compensatory activation of adhesion pathways (integrins, NECTIN1/2, NOTCH) (Document-49649). Functionally, TC-ART impaired wound restitution, reduced mucin coverage, and altered cytokines. In NAM systems, organ-on-chip (OC) models demonstrated more controlled immune phenotypes compared to non-continuous (NC) systems. OC exhibited a quiescent baseline with pronounced transcriptional responses in stress, apoptotic, and transporter pathways (AASS, ABCC1/2 upregulation; ABCB1/4 downregulation; BIRC3 upregulated in OC but downregulated in NC), while selective, moderate cytokine responses with limited increases in inflammatory, chemokine, and growth-factor mediators observed in OC compared to broader dose-dependent fluctuations in NC models.

Conclusions: These findings suggest a biphasic response to TC-ART, with mid-dose inducing adaptive responses and higher exposures causing immune suppression and impaired repair pathways. Our preliminary evidence supports that non-animal models reflect key in vivo outcomes.

Keywords: Multi-omics, Organ-on-chip (OoC) / New Approach Methodologies (NAMs), Tri-combination antiretroviral therapy (TC-ART), Gut microbiome, Sex-specific effects

Poster Number: 7

Integrating transporter data and machine learning approaches to predict drug transport across the placental barrier

Tong Wang[1], Xuelian Jia[1], Lauren M Aleksunes[2], and Hao Zhu[1]

1. Center for Biomedical Informatics and Genomics, Tulane University, New Orleans, LA, USA
2. Department of Pharmacology and Toxicology, Ernest Mario School of Pharmacy, Rutgers University, Piscataway, NJ, USA

Background: The placental barrier is crucial for protecting the fetus from damage caused by medications and toxicants. Drug permeability across the placental barrier is dictated in part by the regulation and function of both uptake and efflux transporters. Animal models for assessing drug placental permeability are expensive, time-consuming, have limited cross-species relevance, and often raise ethical concerns regarding testing in pregnant people. We aimed to integrate drug transporter data with innovative machine learning (ML) approaches to predict drug permeability across the human placental barrier.

Methods: The target activity was the human cord blood-to-maternal blood concentration ratio (CM), which directly reflects overall placental permeability. The transporter data were initially used as descriptors to develop several ML models through a 5-fold cross-validation procedure. Subsequently, the transporter data were used to construct an interpretable neural network (INN) to predict CM and reveal potential routes for drug transport. The neural network was designed to mimic transporters located on multiple membranes (apical and basolateral syncytiotrophoblasts and fetal endothelial cells) within the human placental barrier.

Results: The resulted ML models showed good performance and could predict the drug CM values with coefficients of determination ranging from 0.612 to 0.664. The resulting INN was able to predict CM ratios by activating neurons that represented specific transporters involved in the drug transport routes across the relevant cellular membranes within the placental barrier.

Conclusions: This study provides a novel computational modeling framework that predicts drug placental permeability based on placental transporter functions and regulations, which can be widely applied to assess the risk of drug and toxicant exposures to the fetus during pregnancy. Supported by UC2HD113039 and P30ES005022.

Keywords: Placental barrier, Drug permeability, Transporter, Machine learning, Human cord blood-to-maternal blood concentration ratio

Poster Number: 8

Modeling Female Reproductive Toxicity of Endocrine Disruptor Mixtures: Insights from Targeted Chemicals

Xuelian Jia[1], Tingjie Zhan[2], Congcong Zhang[2], Shuo Xiao[2], Hao Zhu[1,3]

1. Center for Biomedical Informatics and Genomics, Tulane University, New Orleans, LA, USA
2. Department of Pharmacology and Toxicology, Rutgers University, Piscataway, NJ, USA
3. Department of Chemistry & Biochemistry, Rowan University, Glassboro, NJ, USA

Background: The ovaries are central to female reproduction as sites of hormone synthesis, maturation and ovulation of oocytes within follicles, the functional unit of the ovary. Follicle development, steroidogenesis, and ovulation are highly susceptible to endocrine-disrupting chemicals (EDCs). Most of previous research focuses on a single compound; however, humans are exposed to complex EDC mixtures which may target multiple signaling transduction pathways and their interactions. This study applied machine learning frameworks to evaluate the effects of real-world waterborne EDC mixtures on ovarian follicular development and ovulation, and to identify potential driving chemicals and toxicity mechanisms.

Methods: Eight water samples were collected from the Delaware, Passaic, and Raritan Rivers in New Jersey and tested for their effects on ovarian follicular outcomes. Concentrations of over 70 organic chemicals, including PFAS, herbicides, pharmaceuticals, phthalates, and other contaminants, were quantified. Biological activity data were retrieved from PubChem, and assays representing molecular initiating events (MIEs) and key events (KEs) relevant to ovarian follicular outcomes were selected. Machine learning classification models were trained to predict untested chemical activities, and regression models were developed to predict mixture-level ovarian disruption, by integrating bioactivity profiles, exposure concentrations, and chemical descriptors. Model interpretation was used to identify mixture-specific toxicity-driving chemicals and mechanisms.

Results: Chemical composition and concentrations varied across rivers, with Passaic River mixtures showing the greatest chemical diversity and strongest inhibitory effects on hormone secretion, follicle growth and ovulation. Machine learning models were developed for assays including nuclear receptor signaling (e.g., ER, AR, PPAR), mitochondrial dysfunction, p53 activation, and oxidative stress, representing plausible MIEs and KEs, with most models achieving balanced accuracy above 0.7. Regression models to predict different ovarian follicular outcomes achieved R2 of 0.6-0.8. Key driving chemicals such as dibutyl phthalate, bisphenol S, and triclosan were found at elevated concentrations in Passaic River samples and predicted to modulate multiple MIEs and KEs.

Conclusions: This study demonstrates the utility of machine learning–based mixture modeling for assessing female reproductive toxicity of complex environmental EDC mixtures. The integration of multimodal data enables accurate prediction and mechanistic interpretation of how co-occurring environmental chemicals may jointly contribute to adverse ovarian outcomes.

Keywords: Mixtures, Machine learning, Endocrine-disrupting chemicals, Female reproductive toxicity

Poster Number: 9

Functionalizing Chitosan Nanoparticles against A. hydrophila in catfish, a Potential Antibiotic Alternative

Sonya W. Reed [1], Zachariah McGowan[2], Qinglong Jiang[4], Fumiya Watanabe[5], and Grace Ramena[1]*

1. Department of Aquaculture and Fisheries
2. Department of Agriculture
3. Department of Computer Science
4. Department of Chemistry and Physics, University of Arkansas at Pine Bluff
5. Center for Integrative Nanotechnology Sciences, University of Arkansas at Little Rock.

Aeromonas hydrophila causes Motile Aeromonas Septicemia (MAS) in fish, leading to significant economic losses and substantial global financial impact. The inefficacy of vaccines and the rise of antibiotic resistance call for alternative treatments to combat A. hydrophila infections in fish. Plant extracts contain bioactive compounds, including phenols, flavonoids, and carotenoids, which possess antioxidative, antifungal, bactericidal, and bacteriostatic properties. Our preliminary data show that clove (Syzygium aromaticum) extracts, particularly eugenol, exhibit strong antimicrobial effects against A. hydrophila, with no cytotoxic, genotoxic, or growth-inhibitory effects in channel catfish ovary (CCO) cells. However, its volatility limits practical application. To address this, we developed a strategy involving the functionalization of chitosan–tripolyphosphate (CS-TPP) nanoparticles (CS-TPP-E-NPs) for the delivery of eugenol and clove extract. Nanoparticles, due to their small size and large surface area, have demonstrated notable efficacy in drug delivery in human medicine and are increasingly being explored in fish medicine. Chitosan, a natural polymer with non-cytotoxic, antimicrobial, immunostimulant, and biodegradable properties, is ideal for nanoparticle formulation. We have successfully synthesized chitosan NPs, CS-TPP-E-NPs, and CS-TPP-clove oil NPs, and verified and characterized them using Fourier Transform Infrared Spectroscopy (FTIR), Nuclear Magnetic Resonance Spectroscopy (NMR), and Scanning Electron Microscopy (SEM). FTIR confirmed functionalization, while SEM validated the nanoparticles' morphology, composition, and size.

Keywords: Aeromonas hydrophila, Eugenol, chitosan-TPP nanoparticles

Poster Number: 10

Efficient Summarization of Drug Labeling Texts with ChatGPT

Lan Ying[1], Hong Fang[1], Leihong Wu[1], Ingle Taylor[1], Weida Tong[1], Zhichao Liu[2]

1. National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, US
2. Nonclinical Drug Safety, Boehringer Ingelheim Pharmaceuticals, Inc, Ridgefield, CT 06877, USA

Background: FDA-approved drug labeling is essential for ensuring the safe and effective use of medications by healthcare providers and patients, but these documents are often lengthy and complex, containing 17 major sections and multiple subsections. To improve usability, the 2006 Physician Labeling Rule (PLR) in 21 CFR required a half-page "Highlights of Prescribing Information" that summarizes 9 major labeling sections to support efficient retrieval and use of key data. However, producing high-quality Highlights remains a manual, expert-dependent process that is labor-intensive, time-consuming, and variable across reviewers and over time. As a result, many non-PLR labeling documents approved before 2001 still lack Highlights.

Methods: This study examines the potential of large language models (LLMs), such as ChatGPT, to automate the generation of Highlights and improve efficiency, scalability, and consistency in labeling summarization. Specifically, a dataset of 1,730 PLR-formatted labeling documents was compiled and human-authored Highlights of Prescribing Information across nine major sections was compared with ChatGPT-generated summaries, resulting in more than 14,000 section-level summary pairs.

Results: Results showed that 87.99% of the pairs achieved similarity scores between 0.8 and 1.0, indicating a high level of consistency. The median similarity scores appeared to anti-correlate with the original section text length and followed the pattern: Boxed Warning (0.94) > Warnings and Precautions (0.85) > Adverse Reactions (0.83).

Conclusions: Overall, ChatGPT demonstrated strong alignment with expert summaries, particularly for safety-related content, when using carefully designed prompts based on expertise in AI and labeling domain knowledge. Our findings show that ChatGPT reliably generated concise, expert-like summaries of complex drug labeling at scale. As LLMs continue to advance, integrating them into pharmacovigilance, labeling review, and regulatory workflows may improve efficiency, consistency, and evidence-based decision-making.

Keywords: ChatGPT, text summarization, drug information

Poster Number: 11

Global Proteomic Dynamics Identify Key Regulatory Pathways in Pseudomonas aeruginosa Air–Liquid Interface Biofilms

Kidon Sung[1], Miseon Park[1], Jungwhan Chon[2], Ohgew Kweon[1], Saeed A Khan[1]

1. Division of Microbiology, National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, USA
2. Department of Food Science and Biotechnology, Kangwon National University, Chuncheon-si, South Korea

Background: Pathogenic bacteria can form thick, shear-resistant, three-dimensional wrinkle biofilms at the air–liquid interface (ALI), where simultaneous access to oxygen and nutrients enhances survival and antimicrobial resistance. Although Pseudomonas aeruginosa PA14 forms ALI biofilms, the underlying molecular mechanisms remain unclear.

Methods: We performed time-resolved quantitative proteomic analysis of P. aeruginosa PA14 ALI biofilms using planktonic cells as controls. Samples were collected at 48 h (T1), 72 h (T2), 96 h (T3), 120 h (T4), and 144 h (T5). Differentially expressed proteins were analyzed using Cluster of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG), and expression patterns (EPs) were identified.

Results: Biofilm thickness increased most markedly between 48 h and 72 h, followed by gradual growth. A total of 1,460 (T1), 1,789 (T2), 1,959 (T3), 1,938 (T4), and 1,949 (T5) proteins were differentially expressed. Upregulated proteins remained relatively stable over time, whereas downregulated proteins increased from 965 at T1 to over 1,900 from T3 onward. Functional analyses showed that most upregulated proteins were associated with metabolic pathways. Among 121 identified EPs, 224 proteins involved in amino acid and inorganic ion transport and metabolism (EP14) were consistently upregulated. In contrast, approximately 871 proteins related to transcription, signal transduction, and energy production (EP2) were consistently downregulated. Notably, most biofilm-associated proteins were downregulated; only two—3′,5′-cyclic-AMP phosphodiesterase (Icc) and type VI secretion system protein (HcpB)—were consistently upregulated (3.1–9.7-fold).

Conclusions: This study reveals dynamic proteomic changes during ALI biofilm development, characterized by sustained metabolic activity and widespread downregulation of regulatory and biofilm-associated proteins. These findings provide insights into ALI biofilm formation and may inform strategies for biofilm control.

Keywords: air–liquid interface biofilm, Pseudomonas aeruginosa, quantitative proteomics

Poster Number: 12

Enhancing growth and health of largemouth bass fingerlings with single cell proteins and probiotics in basal diets

Arielle Shelby[1], Fard Karim[1], Annik Segree[1], Humphrey Wanjala[1], Sonya Reed[1], Yathish Ramena[1], Grace Ramena[1]*

1. Department of Aquaculture and Fisheries, University of Arkansas at Pine Bluff, 1200 University Dr, Pine Bluff, AR 71601

Background: The ever increasing need to meet the nutritional needs of a growing population projected to reach nearly 10 billion by 2050. Aquaculture presents a promising solution with Largemouth Bass as a high-value species and an emerging food fish for U.S. aquaculture. Advances in technologies such as recirculating aquaculture systems have improved production efficiency and sustainability. However, the scalability of aquaculture remains limited by high feed costs, reliance on fishmeal-based diets, and the associated environmental concerns. The growth, survival, and health of Largemouth Bass are significantly influenced by diet quality.

Methods: This study aims to evaluate the growth and Health of largemouth bass fingerlings with single cell proteins and probiotics in basal diets. Over 90 days, LMB fingerlings were maintained in a RAS and subjected to four dietary treatments: Basal, yeast, spirulina and probiotics. We evaluated the growth and survival, assessed the target stress response genes, and evaluated the gut microbiome and did a histopathological analysis of the distal gut.

Results: The incorporation of probiotics and yeast into LMB diets significantly improved survival and gut health, while growth and feed utilization remained consistent across all diets. Better Cumulative weight gain, better survival in Yeast and Probiotic fed LMB, Yeast had immunostimulant effect, better intestinal homeostasis and reduced stress. We observe an abundance of good microbiome (probiotic) and an increased diversity in Yeast and Spirulina fed LMB.

Conclusions: However, an increased concentration of dietary supplements to achieve optimal specific growth rate for sustainable LMB production needs to be assessed.

Keywords: Largemouth bass, Aquaculture, Supplementary diets, Probiotics, Spirulina

Poster Number: 13

A Web Platform for Curating and Analyzing Single-Cell Cancer Transcriptomes

Jialu Ma[1,2], Mary Yang[1], John Talburt[2], Alex Lee[3]

1. Joint Bioinformatics Graduate Program, University of Arkansas at Little Rock and University of Arkansas for Medical Sciences
2. Department of Information Science, University of Arkansas at Little Rock
3. Department of Biology, University of Arkansas at Little Rock

Background: Single-cell RNA sequencing (scRNA-seq) provides fine-grained resolution of cellular heterogeneity and has become indispensable for advancing cancer research. However, public scRNA-seq datasets remain fragmented and inconsistently annotated, limiting their potential for comparative and translational studies.

Methods: We developed a web platform that systematically retrieves and curates human cancer–related scRNA-seq datasets from the Gene Expression Omnibus (GEO). Our large language model–driven pipeline identified high-confidence cancer single-cell transcriptomic datasets using cancer type and primary site keywords defined in the OncoTree taxonomy. Curated datasets were stored in MongoDB and organized in the portal through a FastAPI backend and React/TypeScript frontend. The portal provides an interactive search interface with keyword and multi-field filtering across cancer types, sample sizes, and publication windows. It also integrates preprocessing modules (quality control, normalization, and metadata harmonization), cell clustering, differential expression analysis, and pathway enrichment analysis.

Results: Our screening pipeline identified over 2000 cancer scRNA-seq datasets after filtering. The portal supports rapid dataset discovery with multi-field filters and is periodically synchronized with new GEO entries. Runtime benchmarking showed that preprocessing and visualization performance depend on dataset size and input format, with .h5ad consistently achieving the fastest performance. Differential expression and pathway analyses were completed within seconds to minutes across the tested datasets.

Conclusions: This web platform provides a user-friendly resource for accessing and analyzing cancer-related scRNA-seq datasets. By combining large-scale curation with integrated analysis workflows, it lowers the barrier to discovery and integration of cancer scRNA-seq resources, thereby facilitating research and clinical applications.

Keywords: single-cell RNA sequencing, cancer transcriptomics, data curation and analysis platform

Poster Number: 14

Detecting Mentions of Substance Use in Clinical Notes

Nahiyan Bin Noor[1], Jonathan P Bona[1]

1. Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, United States

Background: Substance use disorders represent a significant public health challenge, yet identifying drug mentions in electronic health records remains challenging. While Natural language processing (NLP) and large language models (LLMs) offer promising approaches, their effectiveness for multi-label drug mention classification in clinical text remains underexplored; the task is specifically complicated by the presence of multiple different drugs within a single note or section.

Methods: We conducted a comprehensive evaluation of machine learning approaches for multi-label drug detection using the MIMIC-Ext-DrugDetection dataset [1], which contains 8,053 annotated clinical text samples (split into 1,610 samples for training/validation of traditional and deep learning models and a 6,443-sample test split for LLMs) representing eight drug use categories. We compared traditional machine learning models (Logistic Regression, Support Vector Machine etc), deep learning architectures (Bidirectional LSTM, ClinicalBERT), and large language models (Meditron3-8B, Llama-3.1-8B-Instruct) evaluated under zero-shot, few-shot, and fine-tuned configurations.

Results: Fine-tuned large language models achieved the highest performance, with Meditron3-8B attaining a micro F1 of 0.952, macro F1 of 0.812 on the 6,443-sample test split. Among traditional machine learning models evaluated on the 804-sample test set, Linear SVM achieved the best average F1 of 0.799, followed by Logistic Regression (0.720). Among deep learning architectures, ClinicalBERT achieved an F1 of 0.853 on validation, while per-label BiLSTM models reached 0.604. All models struggled with minority classes, particularly prescription opioid misuse (F1 ranging from 0.00-0.53 across non-LLM methods, and 0.108–0.214 for fine-tuned LLMs).

Conclusions: Parameter-efficient fine-tuning of large language models using Quantized Low-Rank Adaptation (QLoRA) substantially outperforms machine learning and deep learning approaches for multi-label drug detection. The medical domain-adapted Meditron3-8B model demonstrated advantages in few-shot learning scenarios, while fine-tuning effectively bridged the performance gap with general-purpose models.

References [1] Harel-Canada, F., Peng, N., Goodman, D., Romero, R., Nguyen, A., Moghanian, B., & Salimian, A. (2025). MIMIC-Ext-DrugDetection (version 1.0.0). PhysioNet. RRID:SCR_007345. https://doi.org/10.13026/0kyx-r485

Keywords: substance use detection; natural language processing; multi-label classification; large language models; clinical informatics; electronic health records; MIMIC-IV

Poster Number: 15

RNA-seq Analysis to Decipher the Interactive Effects of Reactive Oxygen Species Signaling and Lipid Metabolism in Plants

Vanshika Jindal[1], Jiamei Li[1], Fiona Goggin[1]

1. Department of Entomology and Plant Pathology, University of Arkansas, Fayetteville, AR, US

This study utilized RNAseq analysis in the model plant Arabidopsis thaliana to investigate the interactive effects of loci that, together, modulate plant growth and defense, and that, respectively, influence fatty acid metabolism and reactive oxygen species (ROS) signaling in the chloroplast. Fatty Acid Desaturase 7 (FAD7) encodes a desaturase that modulates the desaturation levels of lipids in the chloroplast membranes. EXECUTER1 (EX1) encodes a receptor for singlet oxygen, a powerful ROS generated in the chloroplast, and mediates retrograde signaling between the chloroplast and nucleus. Phenotypic analysis of Arabidopsis null mutants with defects in FAD7 or EX1 and its homolog EX2 suggest that FAD7 and the EX signaling pathway have interactive effects on plant morphology, flowering time, and aphid resistance. We hypothesize that levels of FAD7 activity may influence transcriptional programs through the EX signaling pathway, and to test this hypothesis, we have used RNAseq analysis to compare the foliar transcriptome of four near-isogenic lines with wild-type or mutated FAD7 and/or EX1/EX2. Here we present the results of our transcriptome analysis, and compare two different approaches for exploring the interactive effects of our loci of interest: standard pairwise comparisons between genotypes versus a full factorial analysis. This study will provide methodological insights for RNAseq data analysis as well as biological insights into the interactions between lipid metabolism and ROS signaling in plants.

Keywords: RNAseq analysis, gene expression, ROS signaling, lipids

Poster Number: 16

Modeling the Effect of Menopausal Transition on Heart Failure Risk Using Electronic Health Record Data, a study with All of Us database

Buddhika Jayawardana[1], Dong Wang[1]

1. Division of Bioinformatics and Biostatistics, National Center for Toxicological Research FDA, Jefferson, Arkansas, USA

Background: Menopause is a critical, natural life stage signaling the end of reproductive function, representing a pivotal moment for women's health associated with an increased risk for various diseases. Despite its importance for preventative health, lifestyle adjustments, and monitoring, this transition remains understudied. Using electronic health record (EHR) data holds promise for expanding the data available to study menopause-related health outcomes. However, it is challenging to label menopausal status accurately and incorporate this information into models for disease risk.

Methods: In this presentation, we investigate the relationship between menopausal transition and the risk for heart failure using data from NIH All of Us database. As the exact timing of menopause is usually unknown, we identify a series of marker events from each woman's EHR that are indicative of menopause and use the earliest event to establish a proxy for the onset of the menopausal transition. These marker events include menopausal and postmenopausal disorders, ovarian failure, postmenopausal state related conditions, relevant surgical procedures, and other pertinent conditions specified with ICD9/ICD10 codes. After alignment with age, the onsets of heart failure were identified for each patient in the follow-up time period in the EHR. A Cox proportional-hazards model was constructed for disease onset.

Conclusions: Using this method, we found that the menopausal transition is associated with a significant change in the risk of heart failure. Similar approaches can be applied to other health conditions.

Keywords: survival analysis, menopausal transition, heart failure, women's health, All of Us

Poster Number: 17

A high-quality de novo genome assembly of horsetail milkweed Asclepias subverticillata offers new insights at the plant-insect interface

Mahamad Sayab Miya[1,2], Terrence Sylvester[3,4], Rokeya Akter[1,2], Duane McKenna[3,4], Richard Adams[1,2]

1. Department of Entomology and Plant Pathology, University of Arkansas, Fayetteville, AR, USA
2. Center for Agricultural Data Analytics, University of Arkansas, Fayetteville, AR, USA.
3. Department of Biological Sciences, University of Memphis, Memphis, TN, USA
4. Center for Biodiversity Research, University of Memphis, Memphis, TN, USA

Horsetail milkweed (Asclepias subverticillata) is a toxic herbaceous plant native to North America that serves as a host for many insects of ecological, environmental, and agricultural relevance. Despite its broad importance, no genomic resources exist for this species, limiting insights into the mechanisms underlying its well-known interactions with highly specialized insect herbivores. Here we sequenced, assembled, and annotated the first high-quality de novo genome of A. subverticillata using long-read technology. Genome profiling provided evidence of whole-genome duplications resulting in hexaploidy with relatively high heterozygosity (2.28%). Despite this complexity, our genome assembly spanned 341.49 Mb, distributed across 210 contigs, with an N50 of 14.22 Mb and the largest contig measuring 24.96 Mb. The assembly has a GC content of 33.32% and 97% BUSCO gene representation, indicating high continuity in both the assembly and annotation. High repetitive content (63.44%) was evident in this genome, with a large fraction of interspersed elements and unclassified LTRs (36.72%). We also reconstructed the organelle genomes, with the mitochondrial genome measuring 586,290 bp and containing 40 protein-coding sequences, 23 tRNAs, and 3 rRNA genes. The chloroplast genome is 162,654 bp long and includes 88 protein-coding genes, 8 rRNAs, and 36 tRNAs. Genome annotation recovered 24,939 putative protein-coding genes and 7,150 ncRNAs, with functional predictions for 97.92% of the coding genes. Comparative analysis found the expansions of regulatory and stress/defense-related gene families and contractions of aromatic and terpenoid biosynthetic categories. We identified 56 biosynthetic gene clusters, 82% of which—including all saccharide types—were co-localized with high-confidence CAZymes. The GT1 glycosyltransferase family emerged as the primary tailoring driver, supporting a 'scaffold-and-tailor' strategy for the biosynthesis of defense toxins like cardenolides and cyanogenic glucosides. This genomic resource provides a critical resource for investigating the genomic drivers of long-standing insect-plant interactions in milkweeds and the broader Apocynaceae family.

Keywords: Apocynaceae, cardinolides, hexaploidy, insect-plant interaction, poison milkweed

Poster Number: 18

AI-Based Structural Modeling Reveals Specificity in Peptide Receptor Signaling in Plant Immunity

Fiona L. Goggin[1], Kallahan Minor[1], Jiamei Li[1], Leonidas Dagostino[2], Gunvant Patil[2], Eric Enemark[3]

1. Department of Entomology and Plant Pathology, University of Arkansas, Fayetteville, AR, USA
2. Institute of Genomics for Crop Abiotic Stress Tolerance, Texas Tech University, Denton, TX USA
3. Winthrop P. Rockefeller Cancer Institute, University of Arkansas Medical School, Little Rock, AR USA

Peptide-mediated signaling is a central component of plant innate immunity, yet the extent to which immune peptide receptors vary within crop species remains poorly understood. Such intraspecific variation could critically influence the effectiveness of peptide-based disease control strategies. In soybean (Glycine max), plant elicitor peptides (PEPs) are released upon pathogen attack and activate immune responses by binding Plant Elicitor Peptide Receptors (PEPRs) at the cell surface. Exogenous application of synthetic soybean PEPs (GmPEP1–3) reduces nematode infection in the reference cultivar Williams82, highlighting their potential as sustainable crop protectants. However, the success of this approach across diverse germplasm will depend on the presence and functionality of compatible PEPRs.

Here, we combine AI-based protein structural modeling with population-scale genomics to investigate PEP–PEPR interactions and their conservation within soybean. First, we used AlphaFold Multimer to predict binding specificity between GmPEP1–3 and the three putative soybean PEPRs. These predictions indicate that GmPEPR1b is likely the sole receptor for GmPEP3 and can also interact with GmPEP1 and GmPEP2, suggesting a central role for this receptor in soybean immune signaling. Second, we integrated these structural predictions with variant data from a worldwide soybean diversity panel to assess intraspecific receptor variation. Across all three PEPR genes, we identified high impact mutations predicted to impair receptor function, demonstrating that immune receptor diversity exists within species as well as between species. Notably, most damaging variants were absent from modern breeding lines, and GmPEPR1b was highly conserved among U.S. commercial cultivars.

Together, these findings demonstrate the utility of AI driven structural modeling for dissecting immune signaling pathways, reveal intraspecific PEPR variation as an underappreciated source of immune diversity, and support the broad applicability of GmPEP3 for peptide based disease management in soybean.

Keywords: protein modeling, receptor-ligand interactions, plant immunity, peptide signaling

Poster Number: 19

Investigating the Prevalence and Impact of Heavy Metal Resistance Genes in Salmonella enterica: Links to Antibiotic Resistance

Eric Tang[1], Steven L. Foley[2]

1. Little Rock Central High School, Little Rock, AR 72202
2. National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas 72079, USA

Background: Metals are widely used in animal feed for their growth-stimulating and antimicrobial effects; yet, there is potential concern that their use can promote antimicrobial resistance through co-selection. However, the prevalence of these metal resistance genes in Salmonella and their impact on the induction of antimicrobial resistance remain unclear. To aid in this understanding, this study investigated of the prevalence of heavy metal resistance genes (HMRGs) and their comparison with antimicrobial resistance genes (ARGs) in Salmonella enterica strains isolated from various sources, across different locations and time periods.

Methods: Data on stress and AMR genotypes, serovar, source, location, and collection date were retrieved from the NCBI Pathogen Detection Isolate Browser. Isolates from the United States with complete metadata were analyzed using Microsoft Excel and PANDAS (Python Data Analysis Library). Chi-square tests were conducted to assess differences in antimicrobial resistance genes (ARGs) presence between heavy metal resistance genes (HMRGs) -positive and HMRG-negative isolates. Additionally, the co-localization of HMRGs and ARGs on plasmids was examined, and plasmid incompatibility types were assessed.

Results: A total of 86,968 isolates were analyzed in this study. The prevalence of HMRGs varied markedly across different serovars and isolation sources, indicating heterogeneous distribution patterns. Notably, ARGs were detected at significantly higher frequencies in isolates carrying HMRGs, suggesting a potential association between metal and antibiotic resistance determinants. In addition, co-localization of HMRGs and ARGs on plasmids was commonly observed, although no single plasmid incompatibility group was found to be exclusively linked to this co-occurrence.

Conclusions: Overall, these findings support a potential linkage between heavy metal resistance and antimicrobial resistance in in S. enterica, reinforcing concerns about the use of heavy metals in agriculture. The results provide important insights for risk assessment and inform strategies aimed at mitigating AMR as a global public health threat.

Keywords: United States; heavy metal resistance; antimicrobial resistance; plasmids

Poster Number: 20

Lipid oxidation–inhibitory and antioxidant properties in soybean protein: Integrated experimental and bioinformatic approach

Sukanya Poddar[1], Parvej Hasan Jon[1], Kaustav Majumder[2], Caio Canella Vieira[3], Md Mahfuzur Rahman[1]

1. Department of Food Science, University of Arkansas, 2650 N. Young Ave., Fayetteville, AR, 72704
2. Department of Food Science and Technology, University of Nebraska–Lincoln, 1901 N 21st St., Lincoln, NE 68588
3. Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR, 72704

Plant-derived peptides represent promising natural alternatives to synthetic preservatives for controlling lipid oxidation and preventing food spoilage. Soybean protein is a major plant protein source that can generate peptides with free radical-scavenging (FRS) and metal-chelating properties, which can potentially inhibit lipid oxidation in food. This study investigated how soybean variety selection affects targeted peptide yield in hydrolysates, their composition using SE-HPLC and LC–MS/MS, FRS activity, metal chelation activity, and lipid oxidation inhibition activity. Protein isolates from eight soybean cultivars (SP1–SP8) were hydrolyzed with Alcalase (2% E/S, 2 h, pH 8). Degree of hydrolysis ranged from 9.68 to 12.22%, with peptide concentrations of 1.07–1.27 mmol eq NH_2/g. SP3 and SP6 contained the highest proportions of low-molecular-weight peptides (<1 kDa), while SP4, SP5, and SP6 exhibited the strongest Fe^{2+}-chelating potency (EC_{50} = 261–289 µg/mL). Cu^{2+}-chelating activity was substantially lower (~25%), with SP1 showing superior performance. DPPH scavenging was highest in SP1 at 35% at 1 mg/mL. Lipid oxidation assays demonstrated that SP1 and SP4 achieved significantly higher inhibition after 7 days. LC–MS identified abundant di- and tripeptides (Arg–Phe, His–Phe, Glu–Val–Phe, Leu–Gln, Phe–Ile, Thr–Phe) with high predicted FRS scores (0.37–0.48) and metal-chelating scores (0.23–0.31) determined using AnOxPePred. Using the relative abundance of peptides from LC–MS/MS, along with FRS and CHEL scores from AnOxPePred, the 20 most important peptides were identified. His–Phe achieves the highest combined FRS (0.488) and CHEL (0.311) scores in the entire dataset. SP1 and SP4 were enriched in His–Phe, explaining their enhanced bioactivity. These findings established that varietal differences can significantly influence the lipid-oxidation–modulating properties of hydrolysates, providing a basis for selecting suitable varieties for clean-label antioxidant applications in food preservation.

Keywords: Soy protein hydrolysates, free radical scavenging, metal chelating activity, lipid oxidation, AnOxPePred, clean label

Poster Number: 21

Genomic-island cassette architecture enables interpretable prediction of pathogenic Enterococcus cecorum lineages: Cassette2Vec-EC, a structural genomics and machine-learning approach

Rushikesh R. Lagad[1], Shakil Rafi[2], Aranyak Goswami[1*]

1. Department of Animal Science and Center for Agricultural Data Analytics (CADA), University of Arkansas System Division of Agriculture, Fayetteville, AR, USA
2. Data Science Program (DTSC), University of Arkansas, Fayetteville, AR, USA

*Correspondence: Aranyak Goswami; garanyak@uark.edu

Background: Enterococcus cecorum (EC) is an emerging poultry pathogen whose antimicrobial resistance (AMR) and host-associated traits are often carried on genomic islands (GIs). Standard comparative genomics workflows usually reduce genomes to unordered gene inventories and may miss informative neighborhood structure within island-associated modules.

Methods: We developed Cassette2Vec-EC, a structural-genomics approach that transforms annotated genomes into cassette units (local gene neighborhoods anchored by GI context) rather than treating genomes as unordered gene inventories. Using a curated global cohort of 145 genomes (50 pathogenic, 95 commensal), we integrated pangenome context, mobility markers, and AMR annotations into 84 genome-level cassette-summary features, evaluated under locked 5-fold genome-grouped cross-validation to prevent within-genome leakage.

Results: The Random Forest cassette-summary model achieved AUROC 0.918 ± 0.067 under 5-fold genome-grouped cross-validation, outperforming GI-burden (AUROC 0.791) and assembly-quality (AUROC 0.743) baselines. Cassette features remained discriminatory after controlling for assembly fragmentation (≤50-contig subset AUROC 0.827). Cross-project leave-one-BioProject-out validation on a fully held-out BioProject (n = 53) yielded AUROC 0.694, indicating meaningful but limited transportability. SHAP analysis localized predictive signal to GI-anchored modules co-organizing AMR cargo (tet(M)/tet(L); mefA/msrD) with mobility machinery.

Conclusions: These results show that cassette architecture captures biologically meaningful signal beyond bulk island burden and supports its use as an interpretable genomic representation for surveillance-oriented analysis of poultry-associated E. cecorum. Pathogenic E. cecorum lineages are distinguished not by how many genomic islands they carry but by the content and co-organization of those islands; encoding cassette-level co-organization as interpretable features captures that biologically real signal more actionably than bulk inventory summaries alone.

Keywords: mobile genetic elements; antimicrobial resistance; poultry pathogen surveillance; comparative genomics; accessory genome; interpretable machine learning; horizontal gene transfer; genomic epidemiology

Poster Number: 22

FoodOxPep: Integrating feature fusion deep learning and experimental validation for multifunctional lipid oxidation–inhibitory peptide prediction

Parvej Hasan Jon[1], Sukanya Poddar[1], Aranyak Goswami[2], Mahfuzur Rahman[1]

1. Department of Food Science, University of Arkansas, 2650 N. Young Ave., Fayetteville, AR, 72704
2. Department of Animal Science, University of Arkansas, Fayetteville, AR, 72704

Background: Food-derived bioactive peptides with free radical scavenging (FRS) and metal chelating (CHEL) activities are emerging as natural antioxidants for clean-label food preservation, offering an alternative to synthetic additives for controlling lipid oxidation. However, experimental screening remains labor-intensive and resource-demanding. Computational approaches provide a scalable solution, yet most existing models are designed for nutraceutical purposes and focus on a single bioactivity endpoint and are not tailored specifically for food applications. Therefore, this study aimed to develop and validate a dual-output deep learning model to simultaneously predict FRS and CHEL activities from peptide sequences, with a focus on clean-label preservation.

Methods: FoodOxPep was developed using a feature fusion architecture combining a Bidirectional Long Short-Term Memory (BiLSTM) network with physicochemical descriptors, including amino acid composition, dipeptide composition, and AAindex features. The model was trained using an 80/20 stratified split with duplicate removal. For validation, soy protein isolates were hydrolyzed using Alcalase (2% E/S, 2 h), and peptides were identified by LC–MS/MS. A set of 20 di- and tripeptides was encoded and evaluated for FRS and CHEL prediction.

Results: FoodOxPep achieved training and validation accuracies of 92.14% and 90.51%, respectively. For FRS prediction, precision, recall, and F1-scores were 0.91, 0.88, and 0.89, while CHEL prediction achieved 0.82, 0.79, and 0.73. Among identified peptides (Arg–Phe, His–Phe, Glu–Val–Phe, Leu–Gln, Phe–Ile, Thr–Phe), His–Phe (HF) showed the highest predicted multifunctional activity. Soy protein hydrolysates rich in HF exhibited the highest FRS activity through DPPH radical scavenging activity. Furthermore, samples with higher predicted multifunctional peptide abundance demonstrated improved lipid oxidation inhibition.

Conclusions: FoodOxPep is a tool for identifying multifunctional antioxidant peptides that inhibit lipid oxidation, and it can be further developed into a more robust architecture with a larger array of experimental validation.

Keywords: Machine learning, BiLSTM, peptide prediction, antioxidant, lipid oxidation inhibition

Poster Number: 23

Web-Based Visualization of Virulence Gene Clusters in Bacterial Pathogens

Hawulethu Ndlovu[1], Gerald Shimo[1], Praise Fabiyi[2], Vinay Raj[1]

1. Department of Mathematics and Computer Science, University of Arkansas at Pine Bluff, Arkansas, USA
2. Department of Biology, University of Arkansas at Pine Bluff, Arkansas, USA

Background: Campylobacter species, including Campylobacter jejuni and Campylobacter coli, E. coli, and Salmonella species, are among the leading causes of bacterial gastroenteritis worldwide. Understanding their virulence gene networks is critical for improving public health surveillance and guiding intervention strategies. However, the increasing scale, heterogeneity, and complexity of genomic datasets present significant computational challenges in data integration, visualization, and interpretation.

Methods: We developed a scalable, interactive web-based platform to visualize virulence genes and infection mechanisms of Campylobacter and Salmonella species using curated datasets aggregated from publicly available genomic and microbiological databases.. The visualization system leverages modern full-stack web technologies, including Next.js for server-side rendering and routing, React.js for component-based interactivity, and Tailwind CSS for responsive UI design. DataTables.js enables efficient client-side sorting, filtering, and pagination of large genomic datasets, while Plotly.js provides high-performance, interactive data visualizations for gene cluster exploration. The platform architecture emphasizes modularity, performance optimization, and seamless data integration. Finally, deployment was carried out on Vercel cloud infrastructure.

Results: The platform enables real-time interaction with complex virulence gene datasets, supporting interactive filtering and visualization of gene clusters and host-species relationships. Users can explore zoomable, interactive network visualizations alongside structured tabular data and annotations. System performance evaluations demonstrate efficient handling of the datasets and improved usability compared to static visualization approaches.

Conclusions: This work presents a robust, cloud-deployed visualization system that bridges microbiology and computer science by transforming complex genomic data into interactive, interpretable insights. By integrating scalable web technologies with biological data analysis, the platform enhances accessibility for researchers and public health professionals and establishes a foundation for future extensions in computational genomics and infectious disease modelling.

Keywords: data visualization, web platform, computational biology, full-stack development, public health

Poster Number: 24

Analysis of Virulent Genes and Gene Networks of Bacterial Pathogens in Food Animals

Praise Fabiyi[1], Vinay Raj[1]

1. Department of Biology, University of Arkansas at Pine Bluff, Arkansas, USA

Background: Campylobacter, Salmonella, E. coli, and Listeria are among the most common bacterial pathogens affecting food animals and are major causes of foodborne illness in humans worldwide. The major strains, Campylobacter jejuni, Campylobacter coli, and Salmonella Typhi, survive in poultry gut and pigs' intestines without causing obvious disease, although mild enteritis and growth depression have been reported in young or immunocompromised chicks, but lead to infections like diarrhea (sometimes bloody), salmonellolis, abdominal pain, fever, and rarely, complications like Guillain-Barré syndrome, septic arthritis, osteomyelitis, meningitis, and endocarditis in humans upon ingestion of a carrier animal.

Methods: Virulence-associated genes (cadF, ciaB, cdtA, cdtB, cdtC, flaA, virB11, invA, hilA, prgH, etc.) of the most common strains of Campylobacter and Salmonella in food animals were identified and profiled by a systemic review of the literature. The exon and intron structures of the virulent genes obtained from ENSEMBL gene annotation information were analyzed. For phylogenetic analyses, deduced amino acids sequences encoded by the virulent genes from various species were aligned to assess the evolutionary relationships. The co-occurrence patterns and functional roles were assessed using network pathway analysis. The resistance profiles of the Campylobacter strains, adaptation and impacts to hosts, and their functional relevance to human infection were studied.

Results: Preliminary findings indicate a high prevalence of distinct virulent genes expressions in poultry and other food animals' isolates, with distinct gene networks emerging across host species. Specific gene clusters, including cadFciaB-cdtABC in Campylobacter and sipABC in Salmonella, were strongly associated with traits implicated in human gastrointestinal stress and reduced productivity in poultry.

Conclusions: Our study highlights how virulence gene networks in Campylobacter and Salmonella contribute to colonization dynamics and potential disease expression in humans and food animals. Analyzing these genetic interactions provides valuable insight into zoonotic risk and improves our understanding of animal and human health.

Keywords: Bacteria, Pathogen, Virulence, Gene Networks, Annotation

Poster Number: 25

Python Based Feature Ranking for Heart Failure Survival: A Reproducible Analytics Workflow

Taureen Sprinkle[1], Vinay Raj[1]

1. Department of Mathematics & Computer Science, University of Arkansas at Pine Bluff, Pine Bluff, Arkansas, U.S

Background: Accurate identification of clinically meaningful predictors is central to translational data science and reproducible biomedical research. Chicco and Jurman (2020) demonstrated that survival in heart failure patients can be predicted with high accuracy using only two variables—serum creatinine and ejection fraction—identified through classical biostatistical feature ranking methods implemented in R. Reproducing such analyses using open, widely accessible tools is essential for transparent and reproducible computational workflows. This work aims to replicate and extend the feature ranking methodology of Chicco and Jurman using Python-based statistical and machine learning libraries instead of R, used by the authors. The goal is to gain hands-on experience in biostatistical analysis, reproducible research practices, and the interpretation of clinically relevant predictors.

Methods: Using the publicly available heart failure clinical dataset, we implemented a Python workflow incorporating: Correlation analysis (Pearson) to quantify linear associations between clinical variables and survival. Univariate statistical testing (t tests, chi-square tests) to assess group differences between survivors and non-survivors. Model-based feature importance using logistic regression, random forests, and support vector machines from scikit learn. And finally Reproducibility practices, including environment specification, version-controlled Jupyter notebooks, and automated reporting with Python scripts.

Results: Across statistical and machine learning approaches, serum creatinine and ejection fraction consistently emerged as the strongest predictors of mortality, aligning with the findings of the original publication. Other variables contributed minimal additional predictive value. The Python workflow successfully reproduced the feature ranking hierarchy while offering transparent, fully documented computational steps.

Conclusions: This project demonstrates that Python provides an accessible and reproducible platform for learning biostatistical feature ranking methods in clinical datasets.

Keywords: Python, Reproducible, Heart, Workflow, Data

Poster Number: 26

Peptidoglycan Hydrolases (Pghs) as Alternative Antimicrobials for Streptococcus iniae Infections in Fish

Zachariah McGowan[1], K. Bohara[1], A. Segree[1], G. Ramena[1], D. Donovan[2]

1. Department of Aquaculture and Fisheries, University of Arkansas at Pine Bluff, Pine Bluff, AR, 71601
2. Department of Biology, Morgan State University, Baltimore, Maryland, MD 21251

Streptococcus iniae is a gram-positive pathogen that causes streptococcosis in fish, resulting in high mortality and economic losses in aquaculture. The emergence of antibiotic resistance necessitates alternative therapeutics. Phage-derived endolysins, or peptidoglycan hydrolases (PGHs), enzymes that degrade bacterial peptidoglycan layer, are promising antimicrobials. In this study, we appended eleven cell-penetrating peptide (PTD) domains to two PGHs we previously shown to be active against eight S. iniae strains. PTDs are short peptides that facilitate translocation of PGHs across host cell membranes, enabling intracellular bacterial targeting. We tested these PGH-PTD constructs in OmB cells co-cultured with six S. iniae strains at an MOI of 1:100, treated with 10µg of each purified enzyme. PGH-PTDs lysin one: 7, 10, and 11 significantly reduced intracellular S. iniae strain 35Br, whereas PGH-PTD 12 was effective against strain ATCC 29178. PGH-PTD 7 also lysed strains ARS-9860 and MN15Br, and PGH-PTDs 8 and 9 were active against strains YV16 and BZ1. PGHs lacking PTDs showed minimal intracellular activity, underscoring the essential role of PTDs in enabling cellular uptake. These results demonstrate that PGHs, when fused with PTDs, can effectively target intracellular S. iniae and hold potential as environmentally friendly alternatives to antibiotics for controlling streptococcosis in aquaculture.

Keywords: Streptococcus iniae, Peptidoglycan hydrolases, PGH-PTD, Aquaculture, Antimicrobial alternatives

Poster Number: 27

The Importance of Dataset Recency on Drug-drug Interaction Prediction

Alaa Marwan Abu Taha[1], Grover P Miller[2]

1. Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR
2. Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, AR

Adverse drug reactions caused by drug-drug interactions (DDIs) remain a major challenge to patient safety, particularly in the context of increasing polypharmacy. Despite rapid growth in pharmaceutical innovation, many artificial intelligence (AI) models for DDI prediction continue to rely on outdated datasets like TWOSIDES, which was last updated in 2019. Accelerated drug approvals during and after the COVID-19 pandemic have widened the gap between contemporary clinical practice and the data used to train these models.

In this study, we constructed an updated DDI dataset by parsing FDA Structured Product Labeling files from DailyMed (2026 release) using natural language processing. The resulting dataset captures interactions involving both long-standing medications as well as drugs approved after 2019 that are not present in TWOSIDES. We evaluated the performance of this data using a graph-based neural network across five experimental conditions: (1) baseline replication using TWOSIDES , (2) cross-testing a TWOSIDES-trained model on a randomly sampled subset of the 2026 dataset, (3) evaluation restricted to interactions involving unseen post-2019 drugs, (4) retraining on a size-matched subset of the 2026 dataset, and (5) retraining on the full updated dataset.

While the model reproduced reported benchmark performance on TWOSIDES (ACC = 82.5%, AUC = 90.01%), performance declined when applied to unseen interactions (ACC = 77.96%, AUC = 87.34%) and further decreased for post-2019 drugs (ACC = 75.31%, AUC = 84.28%). Retraining on the 2026 dataset substantially improved performance, with the size-matched subset achieving ACC = 85.23% and AUC = 92.01%, and the complete dataset an ACC = 86.84% and AUC = 93.51%.

These findings demonstrate that dataset recency is a critical determinant of DDI prediction performance. Leveraging current FDA drug label data not only restores model generalizability to modern therapies but also surpasses benchmark results, underscoring the need for continuous data updates to support clinically reliable AI-driven drug safety tools.

Keywords: Machine Learning, Drug Interactions, Artificial Intelligence

Poster Number: 28

Assessing Cardiovascular Risks of Prescription Opioids through Mining Real-World Data in FAERS

Li Ma[1], Johnna Berryhill[1,2], Beverly Lyn-Cook[3], Paul Rogers[1], Weigong Ge[1], Huixiao Hong[1], Weida Tong[1], Wen Zou[1]

1. Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. FDA, Jefferson, AR, USA
2. Department of Mathematics and Statistics, Arkansas State University, Jonesboro, AR, USA
3. Division of Biochemistry and Toxicity, National Center for Toxicological Research, U.S. FDA, Jefferson, AR, USA

Background: Prescription opioids are widely prescribed for pain management, but they are increasingly reported to be linked to cardiovascular adverse events. Understanding these risks is critical for informed prescribing of opioid drugs. Current investigations did not provide a thorough and systematic evaluation of cardiovascular risks related to various kinds of prescription opioids in the context of real-world clinical practice. This study was pursued to address knowledge gaps and comprehensively investigated cardiovascular-related risk profiles of various prescription opioids and their relationships.

Methods: A list of prescription opioids was created using the FDA-approved drug list and the function of "getApproximateMatch" in the RxNorm API. The US FDA's Adverse Event Reporting System (FAERS) database (2004 Q1–2024 Q3) was used to retrieve the adverse event reports. The prescription opioid names were normalized by RxNorm using our published approach, and the Preferred Terms of MedDRA were utilized to filter cardiovascular-related adverse events. The R package "openEBGM" was applied to calculate the Empirical Bayes Geometric Mean (EBGM) scores to identify drug-adverse event pairs.

Results: Around 18 million adverse events reports were retrieved and downloaded from FAERS, and 79,085 disproportionate drug-adverse event pairs were identified by EBGM analysis, linking 17 FDA-approved prescription opioids with 277 distinct cardiovascular adverse events. Examining the distributions of these pairs on drugs and adverse events revealed that fentanyl and oxycodone are the top two drugs with reported cardiovascular adverse events, while hypertension, hypotension, cardiac arrest, cardio-respiratory arrest, and myocardial infarction are the top five cardiovascular adverse events reported. Furthermore, the top 10 cardiovascular adverse events for each of the 17 prescription opioids were identified. Network analysis and hierarchical clustering analysis revealed a close association among fentanyl-remifentanil-sufentanyl and hydrocodone-tramadol-morphine based on their drug-adverse event pairs. In addition, the associations of the 17 prescription opioids based on their cardiovascular advert events profiles showed distinct patterns comparing to the previous published study based on all adverse events profiles.

Conclusions: The results are expected to provide important information and knowledge to help FDA drug regulators, physicians, and patients be aware of cardiovascular toxicities associated with certain prescription opioids or combinations of opioids with other prescription drugs, therefore, preventing or

reducing risks of the prescription opioids-associated cardiovascular disease. The findings highlighted the need for targeted risk assessment and monitoring, could support more refined prescribing practices, and contribute to optimizing opioid safety labeling and regulatory decision-making.

Keywords: prescription opioids, cardiovascular adverse events, FAERS, data mining, network analysis

Poster Number: 29

Exploring Sex Differences in Cardiac ICU Admission and Length of Stay Using Multiple Correspondence Analysis

Shivangi Shrimali[1], Beverly Lyn-Cook[2], Dong Wang[1], and Paul Rogers[1]

1. Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. FDA, Jefferson, AR, USA
2. Division of Biochemical Toxicology, National Center for Toxicological Research, U.S. Food & Drug Administration, Jefferson, AR, USA

Background: Sex differences in Intensive care unit (ICU) admission and treatment are reflected in the underrepresentation of women, poor outcomes, and high mortality risk. Historical exclusion from clinical trials has limited evidence for providing sex-specific care. This study aims to examine data for understanding sex differences in ICU admissions, treatments, and outcomes using electronic health record data, addressing critical gaps in care.

Methods: We used the MIMIC-IV database of de-identified health-related data and conducted multiple correspondence analysis. MCA was applied for data exploration, identifying patterns and associations across categorical variables such as ICU type and demographic factors for both heart failure (HF) and myocardial infarction (MI) patients. The focus was to determine whether sex influenced assignment to specialty cardiac ICUs and evaluate discrepancies in critical care delivery for patients suffering from HF and MI.

Results: Women were associated with admission to non-specialty ICUs. MCA uncovered subgroup patterns showing that admission to cardiac ICUs is not uniform across sex, race, insurance, age, or marital status. Women were associated with admission to non-specialty ICUs, whereas men were admitted to specialty cardiac ICUs. There was also a racial difference noted in assignments to specialty ICUs, with minority groups being underrepresented in specialty ICUs. Additionally, the widowed and single were more frequently admitted to non-specialty ICUs. Length of stay was similar across HF and MI cohorts, with nearly half of patients in each group staying 0–50 hours and was associated with age and marital status in HF and with insurance status in MI, but not with sex.

Conclusions: This data exploration illustrates systemic inequities in ICU allocation based on sex, race, and socioeconomic factors. Evidence of sex-based bias in the medical community highlights the need to address these inequalities in treatments and critical care to achieve better outcomes in the future.

Keywords: sex based bias, demographic analysis, multiple correspondence analysis

Poster Number: 30

An AI-Driven Approach to Predicting 24-Hour ICU Mortality By Digitally Processing Vital Signs Using Bidirectional LSTM Algorithms

Paul Rogers[1]*, Shivangi Shrimali[1], Dong Wang[1]

1. Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food & Drug Administration, Jefferson, AR 72079, USA

* Correspondence: paul.rogers@fda.hhs.gov

Background: Patient vital signs in the ICU are traditionally monitored and recorded on an hourly basis. Variation in vital signs is believed to be key in predicting impending patient death or recovery. Although some patients exhibit significant changes in vital signs as death approaches others do not. Within the ICU, there are a number of factors that can influence patient vital signs including medications and treatments along with the condition or injury from which the patients suffers. Providing advance warning of mortality risk to health care providers allows the opportunity for interventions improving the patient's chances of survival. The increasing number of ICU patients within the US make this topic a significant public health issue.

Methods: The ICU data for this study was obtained from the Medical Information Mart for Intensive Care version IV, generated from the Beth Israel Deaconess Medical Center from 2008 through 2019. We selected subjects who were 20 years and older and in one of the 9 ICUs for at least 24-hours but no more than 3 days.The vital signs data, in addition to use of mechanical ventilation, were restructured into an hourly longitudinal format for each eligible patient over their ICU stay. A Bidirectional LSTM AI algorithm processed these vital signs to make a mortality prediction for each patient.

Results: Our implementation of a bidirectional LSTM AI algorithm coupled with highly granular ICU patient data was predictive of mortality risk. The model could be tuned for either a high sensitivity or high positive predictive value.

Conclusions: The results confirm that there is potential in predicting mortality risk within ICU patients utilizing routinely collected vital signs. An AI algorithm embedded with high sensitivity offers the opportunity for early warning to health care providers concerning patient mortality risk.

Keywords: Mortality, ICU, Death, Artificial Intelligence, Long Short-Term Memory

Poster Number: 31

Chromosome-scale genome assembly of the rice stink bug (Hemiptera: Oebalus pugnax) illuminates genome structure and gene family evolution in Pentatomidae

Rokeya Akter[1], Mahamad Sayab Miya[1], Allen Szalanski[2], Duane McKenna[3], Richard Adams[1]

1. Department of Entomology and Plant Pathology, Center for Agricultural Data Analytics, University of Arkansas, Fayetteville, AR 72701, United States
2. Department of Entomology and Plant Pathology, University of Arkansas, Fayetteville, AR 72701, United States
3. Department of Biological Sciences, University of Memphis, Center for Biodiversity Research, University of Memphis, Memphis, TN 38111, USA

The rice stink bug (Oebalus pugnax; RSB) is a major agricultural pest that poses significant threats to rice production throughout the United States. Yet, despite its importance, almost nothing is known about genome structure, function, and evolution in this species. Here, we sequenced, assembled, and annotated the first high-quality reference genome for RSB and conducted comparative analyses with related hemipteran genomes to understand its historical evolutionary context. The assembly spans 826.62 Mb across 211 contigs, with a N50 of 17.25 Mb and the largest contig of 67.15 Mb. Hi-C-based chromosome-scale scaffolding supported six putative autosomes that together cover 94.20% of the genome. Genome-wide GC content was 30.83%, and repeats comprised 47.55% of the genome, consisting largely of interspersed elements, including DNA transposons, LINEs, and LTRs, with evidence of recent expansions. Our annotation pipeline identified 13,175 putative genes, with predicted functions for 97.87% of protein-coding genes. BUSCO analysis indicated 99.30% completeness, supporting the high quality of the assembly and annotation. Comparative analyses revealed rapid expansion and contraction of gene families associated with feeding, host specialization, and pesticide resistance. Together, these resources provide a new framework for future investigations into genome structure, function, and evolution in this economically and ecologically important insect clade.

Keywords: Pentatomidae, insect-plant interactions, agricultural pest genomics, phytophagy, insect adaptation

Poster Number: 32

Whole Exome Sequencing Identified Genetic Signatures of Immune and Olfactory Pathways in AIH and AIH-ALF patients

Mohammad Sujaur Rahman[1], Mary Yang[1], Minjun Chen[2]

1. Information Science, University of Arkansas at Little Rock, Little Rock, Arkansas, USA.
2. Division of Bioinformatics and Biostatistics, U.S. FDA National Center for Toxicological Research, Jefferson, Arkansas, USA.

Background: Autoimmune hepatitis (AIH) and autoimmune-like hepatitis with acute liver failure (AIH-ALF) are severe immune-mediated liver disorders, but their underlying genetic architecture remains incompletely understood. Objective: To identify shared recurrent variants and functionally relevant genes and biological pathways between AIH and AIH-ALF by comparison with 1000 Genomes controls.

Methods: Significant SNPs were identified independently in AIH and AIH-ALF relative to 1000 Genomes controls. Following variant-level quality control, functional annotation, and gene-based recurrence filtering, overlapping significant SNPs were mapped to genes. Genes harboring at least two shared SNPs were retained as prioritized candidates, yielding a nonredundant set of 77 genes for KEGG pathway enrichment analysis.

Results: KEGG analysis revealed significant enrichment of multiple immune-related pathways, including antigen processing and presentation, allograft rejection, graft-versus-host disease, autoimmune thyroid disease, phagosome, hematopoietic cell lineage, Th17 cell differentiation, and B cell receptor signaling. These enrichments were driven primarily by recurrent HLA-region genes, particularly HLA-B, HLA-DRB1, HLA-DRB5, HLA-DQA1, and HLA-DQB1, with additional contributions from MRC1, KLRC3, LILRB1, LILRB3, and LILRA6. Olfactory transduction was also significantly enriched, supported by olfactory receptor genes including OR2T8, OR2L8, OR2T29, OR2T5, and OR1L6, highlighting an additional signal whose biological relevance requires further investigation.

Conclusions: AIH and AIH-ALF share a core immunogenetic architecture centered on HLA- linked antigen presentation and autoimmune pathway activation. The additional olfactory receptor-related enrichment should be interpreted cautiously but may represent an underexplored component of the shared variant landscape. These findings prioritize candidate genes and pathways for future validation in autoimmune liver injury.

Keywords: Autoimmune hepatitis, acute liver failure, shared genetic variants, HLA region, pathway enrichment

Poster Number: 33

Dose and Time-Dependent Disruption of Intestinal Barrier Genes by a Prohibited Antifungal

Akshath Narravula[1], Sangeeta Khare[1], Kuppan Gokulan[1]

1. Division of Microbiology, National Center for Toxicological Research, U.S. Food and Drug Administration, Little Rock, Arkansas, United States of America

Methods: RNA was extracted from intestinal epithelial cells exposed to varying concentrations (3µm, 30µm, 50µm, 200µm) of "Compound X" for 24 or 72 hrs. DNase treated, and reverse transcribed into cDNA using an Invitrogen SuperScript VILO kit. Gene expression was analyzed using the Qiagen RT² Profiler Human Cell Junction Pathway assay, with GAPDH as the normalization control. Relative expressions were calculated using Qiagen GeneGlobe software. Ingenuity Pathway Analysis (IPA) was used to assess pathway-level effects.

Results: "Compound X" induced dose- and time-dependent changes in gene expression and signaling pathways. Key epithelial barrier genes, including CLDN10, NOTCH2, and ITGB2, were significantly altered. Increased concentrations upregulated CLDN10, suggesting disruption of tight junction integrity, while NOTCH2 downregulation indicates impaired epithelial regeneration. Prolonged exposure decreased ITGB2 expression, implicating compromised cytoskeletal structure. IPA revealed significant enrichment of pathways related to cell junction organization, tight junction signaling, and adherens junctions, confirming barrier disruption. Additionally, activation of leukocyte extravasation, granulocyte adhesion, and IL-8 signaling pathways suggests increased inflammatory responses. Alterations in actin cytoskeleton and integrin, further indicate structural remodeling of epithelial cells. These effects were most pronounced at higher concentrations and longer exposure durations.

Conclusions: This study demonstrates that "Compound X" disrupts intestinal epithelial barrier integrity through coordinated effects on junctional, cytoskeletal, and immune signal pathways. These findings provide mechanistic insight into how residual exposure may increase intestinal permeability, inflammation, and susceptibility to infection, highlighting potential public health risks associated with contaminated seafood.

Keywords: RT-qPCR, Genomics, Public-Health

Poster Number: 34

Proteomic Analysis of δ-Tocoflexol-Mediated Radiomitigation: Plasma Protein Biomarkers of Radiation Injury and Recovery in a Mice Model

Rachael Oluwakamiye Abolade[1,2], Stephen A. Shrum[2], Kim Krager[1], Nukhet Aykin-Burns[1], Cesar M. Compadre[1,2]

1. Department of Pharmaceutical Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205
2. Tocol Pharmaceuticals, LLC, Little Rock, AR 72205.

Background: Exposure to high levels of radiation has a wide range of effects on the body, leading to the development of acute radiation syndrome (ARS). δ-Tocoflexol (DTF), a novel vitamin E analog designed in our laboratory, has demonstrated 70% survival in mice when administered 24 hours after lethal total-body irradiation (8.5 Gy). In this study, we performed a comparative proteomic analysis of plasma protein changes induced by tocoflexol in irradiated and unirradiated mice.

Methods: Male C57BL/6J mice (8–10 weeks) received 8.5 Gy total-body irradiation and were treated with DTF (200 mg/kg, oral/subcutaneous) or vehicle 24 hours post-irradiation. Plasma was collected 10 days post-irradiation. After depletion of abundant proteins with the High Select Top 14 resin, tryptic peptides were analyzed by data-independent acquisition (DIA) mass spectrometry on an Orbitrap Exploris instrument. Spectronaut v20.0 was used to analyze the data against the UniProt Mus musculus database. Statistical comparisons using limma/eBayes identified differentially expressed proteins (DEPs; FDR-adjusted $p < 0.05$, fold change > 2). Ingenuity Pathway Analysis (IPA) was used to map functional networks and canonical pathways.

Results: Four group comparisons were evaluated across 40 mice. Irradiation (Vehicle_IR vs Vehicle) altered the expression of 73 proteins, with notable increases in acute-phase proteins (SAA1, SAA2, haptoglobin, PTX3, hemopexin, tenascin), forming the molecular signature of severe radiation injury. DTF treatment in irradiated mice (Tocoflexol_IR vs Vehicle_IR) reduced DEPs to only 11, demonstrating broad attenuation of the radiation-induced proteomic signature. Cathepsin E, the most severely suppressed protein under irradiation (logFC: −4.58), was significantly restored by DTF (logFC: +1.10, padj: 0.003), with a 5.68 log-fold reversal, representing the most drastic proteomic reversal in the dataset. IPA network and pathway analysis revealed modulation of Acute Phase Response Signaling, LXR/RXR Activation, Cellular Senescence, and DHCR24/cholesterol biosynthesis pathways. DTF alone (Tocoflexol vs Vehicle) produced modest changes (32 DEPs), suggesting that DTF primarily alters protein expression in response to radiation stress rather than under normal conditions.

Conclusions: DTF promotes recovery from radiation injury by reversing radiation-induced changes in protein expression, activating innate immune responses, regulating inflammation and metabolism, and restoring immune cell function. Cathepsin E emerges as a novel molecular target and biomarker of DTF-mediated immune restoration. These proteomic signatures provide mechanistic evidence of its

therapeutic benefit, supporting its development as a radiation medical countermeasure under the FDA Animal Rule.

Keywords: Keywords: proteomics, δ-tocoflexol, radiomitigation, acute radiation syndrome, serum amyloid A, Cathepsin E, biomarkers, vitamin E analogs, tocotrienols

Poster Number: 35

An Integrative Gene Expression Analysis Linking Endocrine Disrupting Chemicals to Prostate Cancer

Md Ali Imam Razu[1,3], Hailemariam Assress[2], Renny Lan[2], Ping-Ching Hsu[1]

1. Department of Environmental Health Sciences, College of Public Health, University of Arkansas for Medical Sciences, Little Rock, AR, USA
2. Arkansas Children's Nutrition Center, Little Rock, AR, USA
3. Bioinformatics, University of Arkansas at Little Rock, Little Rock, AR, USA

Background: Endocrine-disrupting chemicals (EDCs) are widespread environmental exposures that interfere with normal endocrine signaling and further influence gene expression, disrupt enzymatic activity, and impair immune regulation, thereby contributing to hormone-related diseases, including cancer. Prostate cancer, a leading cause of cancer-related mortality, is particularly susceptible to hormone-disrupting chemicals such as bisphenol A, phthalates, parabens, and per- and polyfluoroalkyl substances. However, the molecular mechanisms underlying these associations remain poorly understood.

Methods: We investigated whether genes responsive to EDCs exhibit consistent dysregulation in prostate cancer. A curated list of 1043 EDCs was obtained from the DEDuCT database and mapped to gene interactions using the Comparative Toxicogenomics Database (CTD). Through this process, a subset of EDCs was identified relevant to prostate cancer and linked to 673 associated genes. Expression of these genes was analyzed using bulk RNA-sequencing data from 52 paired tumor–normal samples in The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA-PRAD) cohort. Differential expression was performed using DESeq2, followed by functional enrichment across GO-BP, KEGG, and Reactome databases.

Results: Of the 673 testable EDC-associated genes, 198 were significantly dysregulated (89 upregulated, 109 downregulated; FDR < 0.05, $|\log_2 FC| \geq 0.585$). Pathway analysis revealed a consistent dual pattern: downregulated genes were strongly enriched in xenobiotic metabolism, cytochrome P450, and chemical carcinogenesis pathways, while upregulated genes converged on the KEGG prostate cancer pathway and nuclear receptor/estrogen signaling.

Conclusions: This study demonstrates a practical, hypothesis-driven framework linking environmental EDC exposure to prostate cancer biology, revealing suppressed detoxification alongside activated hormone-driven oncogenic signaling. These insights may inform future mechanistic studies and prevention strategies.

Keywords: Endocrine-disrupting chemicals, Prostate cancer, Differential gene expression, Protein-protein interaction network, Pathway enrichment

Poster Number: 36

Single-Sample Prediction of Average Drug Concentration in ETI Therapy Using Supervised Machine Learning

Saly Abouelenein[1], Ashritha Chalamalla[1,2], E.P.Acosta[3] , K.J. Ramos[4], Jennifer S. Guimbellot*[1,2]

1. Section of Pulmonary and Sleep Medicine, Department of Pediatrics, University of Arkansas for Medical Sciences
2. Arkansas Children's Research Institute
3. Department of Pharmacology and Toxicology, University of Alabama at Birmingham, Birmingham, AL
4. Department of Medicine, University of Washington, Seattle, WA

Background: Elexacaftor (ELX), tezacaftor (TEZ), and ivacaftor (IVA) are used together to treat cystic fibrosis (CF). To support real-world exposure-response studies, including evaluation of side effects and precision therapy, investigators need practical tools to estimate pharmacokinetic exposure from limited sampling. One useful exposure measure is the average drug concentration over time (Cavg), but collecting full pharmacokinetic profiles is often not feasible. This study evaluated whether Cavg can be accurately predicted from a single blood sample collected at any time within 12 hours after dosing. In addition, we examined whether drug concentrations at missing time points could be estimated from observed concentrations at other time points. These approaches may simplify pharmacokinetic analyses and improve use of sparse or incomplete real-world datasets.

Methods: Data from 26 individuals with CF, including both transplant and non-transplant participants, were analyzed. Plasma concentrations of ELX, TEZ, and IVA were measured at nine time points over 12 hours after dosing (0 to 12 h). Cavg was calculated separately for each drug. Supervised learning with leave-one-out cross-validation (LOOCV) was used to evaluate predictive performance. First, separate models were developed for each drug to predict Cavg from a single time point. Next, additional models were used to predict missing concentrations at 3, 7, 9, and 11 hours from the remaining measured time points. Performance was assessed using R^2, mean prediction error (bias), and precision.

Results: Cavg could be predicted well from one blood sample at selected time points. For ELX and IVA, the best performance was observed at 6 hours after dosing, with R^2 values of 0.91 and 0.92, respectively. TEZ performed best at 8 hours (R^2 = 0.87). Missing time-point concentrations were also predicted accurately, with R^2 values ranging from 0.93 to 0.98.

Keywords: cystic fibrosis; CFTR modulators; elexacaftor; tezacaftor; ivacaftor; pharmacokinetics; sparse sampling; predictive modeling; therapeutic drug monitoring

Poster Number: 37

Exploring Artificial Intelligence Usage Patterns Among Undergraduate Students at the University of Arkansas

James Roddy[1], Samantha Robinson[2]

1. Mathematics Department, University of Arkansas, Fayetteville, AR USA
2. Human Environmental Sciences, University of Arkansas, Fayetteville, AR USA

The rapid integration of artificial intelligence (AI) tools into higher education has transformed how undergraduate students approach learning, productivity, and problem-solving. This study investigates AI usage patterns among undergraduate students at the University of Arkansas, with particular emphasis on frequency of use, academic applications, perceived benefits, and ethical considerations. A cross-sectional survey was administered to a diverse sample of undergraduate students across multiple disciplines. The instrument included questions on demographic characteristics, familiarity with AI technologies (e.g., generative AI platforms), frequency and context of use, and attitudes toward academic integrity and institutional policies.

Preliminary findings indicate that a substantial proportion of students regularly utilize AI tools to assist with coursework, including writing support, coding, data analysis, and exam preparation. Students reported increased efficiency and enhanced conceptual understanding as primary benefits. However, notable variation exists across academic disciplines, with STEM and business majors reporting higher utilization rates compared to humanities students. Despite widespread adoption, concerns regarding overreliance, accuracy of AI-generated outputs, and unclear boundaries of acceptable use were frequently expressed.

Additionally, the survey reveals a gap between student usage and institutional guidance, suggesting the need for clearer policies and pedagogical frameworks. Many respondents indicated uncertainty about what constitutes permissible AI assistance, highlighting an opportunity for universities to provide structured guidance and integrate AI literacy into curricula.

This study contributes to the growing body of literature on AI in higher education by providing institution-specific insights and identifying key trends in student behavior. The findings underscore the importance of balancing innovation with academic integrity and suggest actionable recommendations for educators and administrators aiming to effectively incorporate AI into undergraduate education.

Keywords: Artificial Intelligence, Undergraduate Education, Student Behavior, Academic Integrity, Higher Education Technology

Poster Number: 38

Potential role of FAM26F in tumor immune surveillance through Immune synapse formation: Structural and functional similarity to tetraspanin proteins CD9 and CD81

Rabbia Mahdi[1], Samia T Mahmud[1], Adebayo Adetunji[1], Jeffery Osborn[1]

1. Arkansas Colleges of Health Education, Fort Smith, AR

Background: Tumor immune surveillance relies on coordinated communication between immune cells and the tumor microenvironment, mediated by membrane-associated proteins that organize and regulate signaling at the immunological synapse. Tetraspanins such as CD9 and CD81 have gained attention for their roles in immune modulation and cancer biology. FAM26F is a recently identified protein implicated in innate immune activation; however, its structural classification and role in tumor surveillance remain poorly understood.

Hypothesis: FAM26F shares structural and membrane-organizing features with canonical tetraspanins CD9 and CD81, enabling a functional role in immune cell–tumor interactions and surveillance.

Methods: An integrative in silico approach was employed, including sequence analysis, transmembrane topology prediction, and AlphaFold2-based structural modeling. Structural comparisons and visualization were conducted using PyMOL to assess homology with established tetraspanins. Predicted membrane topology, extracellular loop architecture, and disulfide bonding patterns were evaluated alongside protein interaction and co localization predictions.

Results: AlphaFold2 structural modeling and PyMOL visualization demonstrate that FAM26F shares key structural features with CD9 and CD81, including four transmembrane domains and disulfide-stabilized extracellular loops. Notably, the large extracellular (EC2) domain of FAM26F mirrors that of CD9 and CD81, which are known to mediate integrin binding and organize signaling at the immunological synapse. Based on these similarities, we propose a model in which FAM26F co localizes with integrin tetraspanin complexes, contributing to membrane microdomain organization and enhancing immune and tumor communication. These findings further suggest that dysregulation of FAM26F may impact tumor progression and facilitate immune evasion.

Conclusions: The observed structural parallels between FAM26F and canonical tetraspanins, combined with its emerging cancer-associated expression patterns, support the hypothesis that FAM26F functions as both a biomarker and a regulator of tumor immune surveillance. This work provides a strong foundation for subsequent mechanistic validation and highlights FAM26F as a potential target for translational and immuno oncology research.

Keywords: Tetraspanin proteins, Immune synapse formation, Fam26F, Immuno oncology

Poster Number: 39

Rapid assessment of clinical severity for salmonellosis cases via protein family domain analysis and machine learning

Aakash Bhattacharyya[1], Sudip Pandey[1], David Ussery[2]

1. University of Arkansas for Medical Sciences, Department of Biomedical Sciences, Little Rock, Arkansas, USA
2. Oklahoma State University, College of Veterinary Sciences, Stillwater, Oklahoma, USA

Salmonella is a common pathogen, infecting more than a million people yearly. Rapid assessment of clinical case severity is essential for improving patient outcomes and optimizing healthcare resources. Advancements in genome sequencing technologies have enabled the analysis of bacterial genomes from many clinical cases, opening up new opportunities for precise and timely diagnosis. This study proposes a genome-based framework for identifying critical Salmonella cases before the onset of critical symptoms and facilitating early medical intervention. By leveraging protein family (Pfam) domains as the representation for genomic data, the complex genetic profiles of Salmonella cases are simplified into interpretable features. The severity levels of cases were investigated through rigorous data analysis, resulting in a set of 70 Pfam domains that could be potentially used as biomarkers. Machine Learning was employed to assess the predictive power of the curated Pfam biomarkers, achieving high accuracy (~93%) in sorting cases into critical, moderate, and mild categories. The results demonstrate the efficacy of the proposed approach. This framework highlights the potential of using bacterial genomic data in clinical decision-making, opening the window for timely personalized interventions for Salmonella infection management.

Keywords: Genomics, Salmonella, Machine Learning, Severity-Prediction

Poster Number: 40

Variable Level Patterns in Diabetes Progression: Insights from CDC Health Indicators

Alexis Whitten[1], Vinay Raj[1]

1. Department of Mathematics and Computer Science, UAPB, Pine Bluff, AR, USA

Background: Diabetes remains one of the most prevalent chronic conditions in the United States and continues to rise alongside increasing rates of obesity, sedentary behavior, hypertension, and socioeconomic disparities. Key health indicators, including body mass index, blood pressure, cholesterol levels, physical activity, smoking status, income, and education, show consistent differences across population groups and tend to shift gradually over time. These cumulative metabolic and social risk factors shape disease progression. Examining how these variables differ among healthy, prediabetic, and diabetic populations provides important context for identifying early warning signs and prevention opportunities.

Methods: This study analyzes the CDC Diabetes Health Indicators Dataset, a high-dimensional dataset (2000–2014) derived from the U.S. Renal Data System and the Behavioral Risk Factor Surveillance System. The dataset includes 35 demographic, clinical, and lifestyle variables and classifies individuals as diabetic, prediabetic, or healthy. The analysis focuses on identifying observable patterns across variables to understand how gradual changes in health behaviors, physiological measures, and social determinants correspond with worsening disease status. Data were organized and visualized using Python and Excel to calculate summary statistics, compare trends across groups, and highlight consistent risk-related trajectories. The emphasis is on pattern recognition rather than predictive modeling that reflects cumulative risk.

Results: Initial findings show steady increases in body mass index, hypertension, cholesterol abnormalities, and physical inactivity as individuals progress from healthy to prediabetic to diabetic classifications. Higher smoking prevalence and reduced healthcare access also align with worsening metabolic outcomes. Lower income and education levels appear more frequently among diabetic groups, underscoring the influence of social and economic conditions on disease patterns.

Conclusions: This analysis demonstrates how a structured review of variable-level trends can reveal meaningful insights into diabetes risk and progression. Recognizing these gradual shifts may support earlier identification of at-risk individuals and inform targeted prevention strategies.

Keywords: Diabetes, Variables, CDC, Python

Poster Number: 41

Safety Concept Annotation Tool: A Web Based Tool for Annotating Clinical Concepts in Individual Case Safety Reports

Brandon Penny[1], Leihong Wu[1], Joshua Xu[1]

1. Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, Jefferson, Arkansas USA

Post-marketing safety surveillance relies heavily on free-text narratives in Individual Case Safety Reports (ICSRs), but large-scale extraction of clinical concepts (i.e., key safety information) remains challenging. A major obstacle is the absence of extensive benchmarking datasets for evaluating novel annotation strategies. To address this challenge, we developed the Safety Concept Annotation Tool, a web application that provides a unified platform for both interactively annotating narratives and ingesting annotations generated through automated strategies. This platform not only provides users with a structured workflow for creating expert annotations from these reports, expediting the creation of benchmark datasets, but also enables users to overlay and compare manually curated annotations against those produced by automated annotation strategies.

Keywords: ICSR, annotations, AI, web, interactive