5th Annual Meeting of the Arkansas Bioinformatics Consortium

AR-BIC 2019

We are an Arkansas Collaborative Community in Bioinformatics Research

http://www.arkansasbioinformatics.org/

Bioinformatics in Food and Agriculture

February 25-26, 2019

Fred Smith Conference Center
Jackson T. Stephens Spine and Neurosciences Institute
UAMS, Little Rock, AR

ArkansasResearchAlliance
Coming together to move Arkansas forward.

ARBIC
ARKANSAS BIOINFORMATICS CONSORTIUM

FDA U.S. FOOD & DRUG ADMINISTRATION
NATIONAL CENTER FOR TOXICOLOGICAL RESEARCH

ASU ARKANSAS STATE UNIVERSITY

UALR UNIVERSITY OF ARKANSAS AT LITTLE ROCK

UAMS UNIVERSITY OF ARKANSAS FOR MEDICAL SCIENCES

UNIVERSITY OF ARKANSAS

UNIVERSITY of ARKANSAS AT PINE BLUFF

Conference Sponsors:

ArkansasResearchAlliance
Coming together to move Arkansas forward.

FDA U.S. FOOD & DRUG ADMINISTRATION
NATIONAL CENTER FOR TOXICOLOGICAL RESEARCH

UAMS

Arkansas Biosciences Institute

UALR University of Arkansas at Little Rock

STATE ARKANSAS STATE UNIVERSITY

UAPB University of Arkansas at Pine Bluff

University of Arkansas Research & Economic Development

Table of Contents

About the Arkansas Bioinformatics Consortium (AR-BIC)

Mission:

The Arkansas Bioinformatics Consortium (AR-BIC) is a virtual Arkansas-centric bioinformatics community aimed at developing, leveraging and enhancing state-wide collaboration, thus forming a stable environment available to support the Arkansas-wide research, education, training and entrepreneurial/industrial activities in life sciences-related computing. AR-BIC activities are within the general area of life sciences computing in Arkansas. The goals of AR-BIC are to (1) strengthen Arkansas' ability to compete at national and international levels for research funding, (2) enable and facilitate collaboration in research where synergy is identified, (3) enhance education, training and university curricula, and (4) expand Arkansas economic growth and job opportunities. AR-BIC is founded on the belief that we can be more than the sum of our parts, and that in our unity, we can draw strength from our diversity. Through synergy, a true critical mass of capability can be assembled to take on large challenges in public health.

AR-BIC Governing Board

Institute	Governing Board
Arkansas Research Alliance (ARA)	Jerry B. Adams President/CEO
	Bryan J. Barnhouse Vice President
	Julie LaRue Senior Project Manager
National Center for Toxicological Research (NCTR)	William Slikker, Jr., Ph.D. Director
	Weida Tong, Ph.D. Director, Division of Bioinformatics and Biostatistics
	Shraddha Thakkar, Ph.D. Scientist, Division of Bioinformatics and Biostatistics
Arkansas Economic Development Commission (AEDC)	Tom Chilton Director, Science and Technology
University of Arkansas for Medical Sciences (UAMS)	Lawrence E. Cornett, Ph.D. Vice Chancellor for Research
Arkansas Biosciences Institute (ABI)	Robert McGehee, Jr., Ph.D. Director
University of Arkansas (UA)	Ralph Davis, Ph.D. Associate Vice Provost for Research and Economic Development
University of Arkansas at Little Rock (UALR)	Abhijit Bhattacharyya, Ph.D. Interim Vice Provost for Research and Dean of Graduate School
University of Arkansas at Pine Bluff (UAPB)	Mansour Mortazavi, Ph.D. Vice Chancellor for Research and Innovation
Arkansas State University (A-State)	Andrew Sustich, Ph.D. Associate Vice Chancellor of Research
	Thomas Risch, Ph.D. Professor of Animal Ecology

AR-BIC 2019 Steering Committee

Institute	Steering Committee
Arkansas Research Alliance (ARA)	Jerry B. Adams
	Art Norris
	Bryan Barnhouse, MPA
	Julie LaRue
National Center for Toxicological Research (NCTR)	Weida Tong, Ph.D.
	Steven Foley, Ph.D.
	Shraddha Thakkar, Ph.D.
The University of Arkansas (UA)	Doug Rhoads, Ph.D.
	Andy Pereira, Ph.D.
	Michael Thomsen, Ph.D.
	Steven Ricke, Ph.D.
	Young Min Kwon, Ph.D.
University of Arkansas for Medical Sciences (UAMS)	Michael S. Robeson, Ph.D.
University of Arkansas at Little Rock (UALR)	Jerry Darsey, Ph.D.
	Mary Yang, Ph.D.
Arkansas State University (A-State)	Argelia Lorence, Ph.D.
	Elizabeth Hood, Ph.D.
University of Arkansas Pine Bluff (UAPB)	Joseph Onyilagha, Ph.D.
	Vinay Raj, Ph.D.

Venue:

Fred Smith Conference Center – 12thFloor
Jackson T. Stephens Spine and Neuroscience Institute
University of Arkansas for Medical Sciences
4301 West Markham
Little Rock, AR
72205

Contact Information:

Conference Logistic:

Julie LaRue, Senior Project Manager
Arkansas Research Alliance
1125 Oak Street
Conway, AR 72032
Email: jlarue@aralliance.org
501-450-7818 (office); 501-269-8935 (cell)

Scientific Liaison:

Weida Tong, Ph.D.
Director, Division of Bioinformatics and Biostatistics
FDA/National Center for Toxicological Research
3900 NCTR Road
Jefferson, AR 72079-9502
E Mail: weida.tong@fda.hhs.gov
Phone: 870-543-7142

Shraddha Thakkar, Ph.D.
Principal Investigator, Division of Bioinformatics and Biostatistics
FDA/National Center for Toxicological Research
3900 NCTR Road
Jefferson, AR 72079-9502
E Mail: shraddha.thakkar@fda.hhs.gov
Phone: 870-543-7068

Conference sponsors and acknowledgement:

* Arkansas Research Alliance (ARA)
* University of Arkansas (UA)
* University of Arkansas at Little Rock (UALR)
* Arkansas State University (ASU)
* Arkansas Biosciences Institute (ABI)
* University of AR for Medical Sciences (UAMS)
*University of Arkansas at Pine Bluff (UAPB)

**Funding for this conference was made possible, in part, by the Food and Drug Administration through grant 1R13FD005304-01, views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the united states government*

Conference Program
Bioinformatics in Food and Agriculture

Day 1: Monday, February 25, 2019- Morning Session	
8:00 am – 5:00 pm	Registration and Poster Set Up *12th floor lobby, Spine Institute*
9:00 am – 12:00 pm	**Data Analytics Workshops – Concurrent Sessions A, B, and C** (*hands on – bring your wifi-ready laptop*) **Session Chair** - Douglas Rhoads, Ph.D., Director of Interdisciplinary Graduate Program in Cell and Molecular Biology, University of Arkansas, Fayetteville, AR
Jerry Adams	**Concurrent Session A: Beginning Hands-on Workshop for R: Getting Started and Basic Analyses** Samantha Robinson, Ph.D., Clinical Assistant Professor, Department of Mathematical Sciences, University of Arkansas, Fayetteville ***LOCATION: Stephens Spine Institute, Fred Smith Conference Center, 12th floor***
Art Norris	**Concurrent Session B: Advanced Hands-on Workshop for R: Practical Hierarchical Bayesian Modeling** Giovanni Petris, Ph.D., Professor and Director of Statistics, University of Arkansas, Fayetteville ***LOCATION: Cancer Institute – Betsy Blass Conference Room – 10th Floor, Room 1013***
Bryan Barnhouse	**Concurrent Session C: CyVerse Cyberinfrastructure for Research and Education in Genomics and Metagenomics** Jason Williams, Education, Outreach, and Training Lead**,** Cold Spring Harbor Laboratory, Cold Spring Harbor, NY ***LOCATION: Institute on Aging – Room 1-180***

NOTE: Lunch served to workshop participants only – served in workshop meeting rooms.

Day 1: Monday, February 25, 2019 - Afternoon/Opening Session *LOCATION: Fred Smith Conference Center, 12TH Floor, Spine Institute*	
1:00 pm – 5:00 pm	**Session 1: Microbiology in the Land of Data** **Session Co-Chairs –** Steven Foley, Ph.D., Deputy Director, Division of Microbiology, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR; Ohgew Kweon, Ph.D., Staff Fellow, Division of Microbiology, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR
1:00 pm – 1:15 pm	**Governor's Welcome Video** **Jerry Adams, Arkansas Research Alliance** **Bill Slikker, Director, National Center for Toxicological Research (NCTR)**

1:15 pm – 2:15 pm	**Keynote 1: GenomeTrakr and GalaxyTrakr – Distributed Sequencing and Bioinformatics for Food Safety** Errol Strain, Ph.D., Supervisory Mathematical Statistician, Division of Microbiology, Center for Food Safety and Applied Nutrition (CFSAN), U.S. FDA, White Oak, MD
2:15 pm – 2:45 pm	**Antimicrobial Resistance Dissemination Among Enteric Bacteria** Steven Foley, Ph.D., Deputy Director, Division of Microbiology, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR
2:45 pm – 3:15 pm	**Food Production and Microbiome Applications** Steven C. Ricke, Ph.D., Professor, Donald "Buddy" Wray Chair in Food Safety, Department of Food Science, University of Arkansas, Fayetteville, AR
3:15 pm – 3:30 pm	***Break***
3:30 pm – 4:00 pm	**How Bioinformatics Impact Plant Movement Across Border Lines** Ioannis Tzanetakis, Ph.D., Faculty Director, Department of Plant Pathology, University of Arkansas, Fayetteville, AR
4:00 pm – 4:30 pm	**Recent Data on the Impact of Veterinary Drug Residues in Food Animal Derived Products on the Human Intestinal Microbiome** Ahn Young-Beom, Ph.D., Staff Fellow, Division of Microbiology, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR
4:30 pm – 5:00 pm	**Keeping Track of It All: Making Microbiome Data Science Scalable, Extensible, and Reproducible** Michael S. Robeson II, Ph.D., Assistant Professor, Department of Biomedical Sciences, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR
Day 1: Monday, February 25 – Evening Session – Reception and Poster Session *LOCATION: 12TH floor lobby, Spine Institute*	
5:00 pm - 7:00 pm	**RECEPTION** **Poster session** – Theme: Bioinformatics Activities in Arkansas

Day 2: Tuesday February 26th, 2019 – Morning Session *LOCATION: Fred Smith Conference Center, 12TH Floor, Spine Institute*	
8:00 am - 12:00 pm	**Session 2: Data Analytics and Machine Learning** **Session Co-Chairs –** Shraddha Thakkar, Ph.D., Principal Investigator, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. FDA, Jefferson, AR; Bryan Barnhouse, Chief Operating Officer, Arkansas Research Alliance, Conway, AR
8:00 am – 9:00 am	**Keynote 2: Big Data and Machine Learning to Predict Chemical Toxicity** Thomas Hartung, Ph.D., Professor, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
9:00 am – 9:30 am	**Text and Gene – Analysis of Big Genomics Data with Text Mining Methods** Weida Tong, Ph.D., Director, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR
9:30 am – 10:00 am	**Machine Learning Methods in PlantCV for Leaf Tracking and More** Noah Fahlgren, Ph.D., Director, Bioinformatics Core, Donald Danforth Plant Science Center, St. Louis, MO
10:00 am – 10:30 am	***Break***
10:30 am – 11:00 am	**From Better Corn to New Plant Paradigms** Dan Berleant, Ph.D., Professor, University of Arkansas at Little Rock, Little Rock, AR
11:00 am – 11:30 am	**Developing an Intelligent System for Species Identification of Food-Contaminating Beetles** Joshua Xu, Ph.D., Division of Bioinformatics and Biostatistics, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR
11:30 am – 12:00 pm	**For the Blockchain/Internet of Things** Chase Rainwater, Ph.D., College of Engineering, Department of Industrial Engineering, University of Arkansas, Fayetteville, AR
12:00 pm – 1:00 pm	***LUNCH*** *– 12th Floor Lobby, Spine Institute*
	Day 2: Tuesday, February 26th, 2019 – Afternoon Session *LOCATION: Fred Smith Conference Center, 12TH Floor, Spine Institute*

1:00 pm – 5:30 pm	**Session 3: Bioinformatics Resources** **Session Co-Chairs –** Weida Tong, Ph.D., Director, Division of Bioinformatics, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR; Argelia Lorence, Ph.D., Professor of Metabolic Engineering, Arkansas State University (A-State), Jonesboro, AR
1:00 pm – 2:00 pm	**Keynote 3: Real World Evidence Drives Innovation in Health Care** Lawrence Lesko, Ph.D., University of Florida, Gainesville, FL
2:00 pm – 2:30 pm	**Database to Store High-Throughput Phenotyping Data** Jack Cothren, Ph.D., Professor of Geography, Director of the Center for Advanced Spatial Technologies, University of Arkansas, Fayetteville, AR
2:30 pm – 3:00 pm	**Agricultural Data – Supplying Information and Data Needs** Eugene Young, Regional Director, National Agricultural Statistics Service, Delta Regional Office, USDA, Little Rock, AR
3:00 pm – 3:30 pm	***Break***
3:30 pm – 4:00 pm	**Animal QTLdb and CorrDB: Helping to Close the Genotype to Phenotype Gap** James M. Reecy, Ph.D., Professor, Associate Vice President for Research, Iowa State University, Ames, IA
4:00 pm – 4:30 pm	**Computer Cyber Security Training and Food Science** Dale R. Thompson, Ph. D., Associate Professor, Associate Department Head for Undergraduate Programs, University of Arkansas, Fayetteville, AR
4:30 pm – 5:00 pm	**Plant Imaging Consortium in Arkansas** Fiona Goggin, Ph.D., Professor, Director Experiment Station, University of Arkansas, Fayetteville, AR
5:00 pm – 5:30 pm	**Ricebase: A Web Resource for Rice Breeding and Genetic Discovery** Jeremy Edwards, Ph.D., ARS Dale Bumpers National Rice Research Center, USDA, Stuttgart, AR
5:30 pm	**Closing Remarks and Poster Awards Announcement and Presentation** Weida Tong, Ph.D., Director, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research (NCTR), U.S. FDA, Jefferson, AR

AR-BIC Governing Board Biographies

Jerry B. Adams
President/CEO,
Arkansas Research Alliance,
Conway, AR

Jerry Adams is the President/CEO of the **Arkansas Research Alliance**, an economic development non-profit modeled on the very successful Georgia Research Alliance. Its primary focus is to leverage university-based job-creating research in Arkansas. The ARA Board of Trustees consists of the five chancellors of the Arkansas research universities and sixteen Arkansas based CEOs (www.aralliance.org).

Jerry retired from **Acxiom Corporation** in October, 2007 after 34 years serving a variety of senior leadership roles and started the Arkansas Research Alliance in April 2008.

Jerry has an extensive history of public service that includes education (both higher education and Pre-K – 12), entrepreneurship/innovation, healthcare and philanthropy.

Jerry has served on advisory boards at the **University of Arkansas, the University of Arkansas for Medical Sciences, University of Arkansas Little Rock, University of Central Arkansas and the University of the South (TN)**. He has also invested board leadership with the **EAST Initiative**, the **STEM Coalition**, **Arkansas Initiative for Math & Science, Arkansas Commitment** and **Arkansas Pre-school Plus.**

He has been involved with much of the knowledge-economy visioning and implementation with **Accelerate Arkansas, Funds for Arkansas' Future, the Arkansas Regional Innovation Hub and VIC Technology Ventures.**

Jerry serves on the board of the **Arkansas Center for Health Improvement** (ACHI) and was the founding board chair for the **Conway Interfaith Clinic**, focused on the medically underserved in Conway, AR.

Each of the last three governors of Arkansas have appointed Jerry to Blue Ribbon Commissions – **Higher Education (Huckabee), Healthcare (Beebe), Computing & Data Analytics (Hutchinson)** – reflecting a strong commitment to the state of Arkansas.

The **Winthrop Rockefeller Foundation** and **WINROCK International** are current board positions along with a long term involvement with the **Arkansas Community Foundation**.

This past year Jerry received the **Humanitarian of the Year award** from **Just Communities of Arkansas** and the **Distinguished Service Award** from the **Conway Chamber of Commerce**.

Jerry is married with two grown sons and four grandchildren. He and his wife, Madelyn, reside in Conway, Arkansas.

Bryan J. Barnhouse , MPA
Vice President,
Arkansas Research Alliance
Conway, AR

Bryan comes to the Arkansas Research Alliance (ARA) as an experienced economic developer. In his previous six years with the Economic Development Alliance for Jefferson County (Arkansas), he directed business recruitment and retention, which included landing a $3.7 billion project that converts natural gas to clean diesel on more than 1,000 acres, and led programs that re-oriented the research and curriculum of higher education institutions to match corporate needs in the area. Prior to that, he managed the promotion of industrial recruitment from Asia and global trade and export development for the Arkansas Economic Development Commission.

Before relocating to Arkansas, he spent four years at the International City/County Management Association in Washington, D.C., coordinating federal business development activities and managing military and technology projects. His educational background includes a Master of Public Administration with an emphasis on Intergovernmental Management and a Bachelor of Arts in International Relations from the University of Southern California.

Bryan is committed to helping advance his adopted state of Arkansas through the creation of economic development opportunities. He is excited to be pursuing this work with ARA by connecting the state's rich scientific and research communities to each other and to the Arkansas business community. He helps professionalize the practice of economic development as a member of the board of directors for Arkansas Economic Developers. And, as a nine-year member and former Chair of the Little Rock Sister Cities Commission, he helps cultivate relationships and partnerships between Little Rock and cities around the world, particularly with Newcastle Upon Tyne, England, for which he serves as the Commission co-liaison.

Bryan and his wife, Jennifer, reside in the vibrant South Main area of downtown Little Rock and enjoy being part of its growth. They, along with their two rescue dogs, love to explore the outdoors of the Natural State.

Julie LaRue
Senior Project Manager
Arkansas Research Alliance

Julie LaRue is the Senior Project Manager for the Arkansas Research Alliance, an economic development non-profit modeled on the very successful Georgia Research Alliance. Its primary focus is to leverage university-based job-creating research in Arkansas. The ARA Board of Trustees consists of the five chancellors of the Arkansas research universities and sixteen Arkansas based CEOs. (www.aralliance.org)

Julie served for the U. S. Senate Committee on Finance from 1981 – 1983 under Senator Bob Dole, then for the Government Affairs division of Baxter Healthcare in Washington, D.C. from 1983 – 1986. From there, she served as Marketing Manager at Arkansas Blue Cross Blue Shield during 1986 – 1999.

Julie was the Executive Director for the Arkansas Community Foundation / Faulkner County from 2008 – 2014, and has been with the Arkansas Research Alliance since September 2014.

Julie and her husband, Wayne, reside in Conway, Arkansas

William Slikker, Jr., Ph.D.
Director
National Center for Toxicological Research
US FDA
Jefferson, AR

Dr. William Slikker, Jr. is the director of FDA's National Center for Toxicological Research (NCTR). He received his Ph.D. in Pharmacology and Toxicology from the University of California at Davis. Dr. Slikker holds adjunct professorships in the Departments of Pediatrics, and Pharmacology/ Toxicology at the University of Arkansas for Medical Sciences. He is currently associate editor for *NeuroToxicology* and for *Experimental Biology and Medicine*. He has served as president of the Academy of Toxicological Sciences, the Teratology Society and the Society of Toxicology. Dr. Slikker has co-authored over 350 publications in the areas of transplacental pharmacokinetics, developmental neurotoxicology, systems biology, and risk assessment.

Weida Tong, Ph.D.
Director
Division of Bioinformatics and Biostatistics
National Center for Toxicological Research
US FDA
Jefferson, AR

Dr. Weida Tong is Director of Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR/FDA). He has served science advisory board for several multi-institutional projects in Europe and USA. He also holds adjunct appointment at several universities. In addition, he is the founder and board chairperson of newly established international MAQC Society. His division at FDA is to develop bioinformatic methodologies and standards to support FDA research and regulation and to advance regulatory science and personalized medicine. The most visible projects from his group are (1) conducting the Microarray and Sequencing Quality Control (MAQC/SEQC) consortium to develop standard analysis protocols and quality control metrics for emerging technologies to support regulatory science and precision medicine; (2) development of liver toxicity knowledge base (LTKB) for drug safety; (3) in silico drug repositioning for the enhanced treatment of rare diseases; and (4) development of various tools such as ArrayTrackTM suite to support FDA review and research on pharmacogenomics. In addition, his group also specializes in molecular modeling and QSARs with specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book chapters.

Shraddha Thakkar, Ph.D.
Division of Bioinformatics and Biostatistics
National Center for Toxicological Research
US FDA
Jefferson, AR

Dr. Thakkar works at FDA's National Center for Toxicological Research. Her research interests are in applying bioinformatics and cheminformatics for study of toxicity and drug development with specific interest in drug-induced liver injury. She has received multiple research and leadership awards regionally and nationally and with FDA. That includes Genentech Innovation in Biotechnology Award from American Association of Pharmaceutical Scientist (AAPS), Margret C. Etter Student lecturer award from American Crystallography Association, and Outstanding Service award from FDA. Dr. Thakkar has adjunct appointments at both University of Arkansas for Medical Sciences and University Arkansas at Little Rock (Assistant Professor). Furthermore, Dr. Thakkar was elected as Board member of the Mid-South Computational Biology and Bioinformatics Society (MCBIOS) in 2014 and served as President for the Society from 2016-2017. She is also the Chair of Pharmacogenomics Group (2018-19) and serves on the Awards Committee at American Association of Pharmaceutical Scientist (AAPS).

Thomas Chilton, J.Ed.
Director
Science and Technology for the
Arkansas Economic Development Commission
Little Rock, AR

Thomas Chilton is the Director of Science and Technology for the Arkansas Economic Development Commission. He has served for 12 years at the AEDC leading the Arkansas statewide commercialization effort under the Accelerate Arkansas banner- driving the research to jobs initiative.

Before coming to AEDC he was the Director Service Sales and Support for North America at Cisco Systems of San Jose, California. He also was employed as an executive at Alltel Corporation for 10 years in various sales and operational positions, with the major accomplishment of overseeing the building of Alltel Enterprise Network (AEN). He was a partner in the consulting firm Telecommunications International with clients in both the private and public sector.

Thomas has a Juris Doctor Degree from the University of Arkansas Little Rock. He has been a member of the Arkansas Bar since 1997.

Lawrence E. Cornett, Ph.D.
Vice Chancellor for Research
University of Arkansas for Medical Sciences
Little Rock, AR

Dr. Cornett is a professor in the Department of Physiology and Biophysics at the University of Arkansas for Medical Sciences and the Vice Chancellor for Research. Dr. Cornett earned a BS in biology from the University of California, Riverside, his Ph.D. in physiology from the University of California, Davis, and was a postdoctoral fellow in reproductive endocrinology and cardiovascular physiology at the University of California, San Francisco. His research interests include the role of 2-adrenergic receptors in mediating airway responsiveness in asthma and hormonal regulation of stress responses at the level of the pituitary gland. In addition, he is the Director of the Arkansas IDeA Network of Biomedical Research Excellence (INBRE), a program funded by the National Institutes of Health to develop biomedical research infrastructure in the state. Dr. Cornett is a member of the American Association of Medical Colleges GRAND Steering Committee, the Arkansas Children's Research Institute Board of Directors, and the EPSCoR/IDeA Foundation Board. Among his many honors, Dr. Cornett received a fellowship from the NIH Fogarty Center and a Research Career Enhancement Award from the American Physiological Society.

Robert McGehee, Jr., Ph.D.
Director and Professor
Arkansas Biosciences Institute
University of Arkansas for Medical Sciences
Little Rock, AR

Robert E. (Bobby) McGehee, Jr., Ph.D., is a Professor of Pediatrics in the UAMS College of Medicine. He holds joint appointments in the Department of Physiology and Biophysics and the Department of Pathology. McGehee also serves as Dean of the UAMS Graduate School and Executive Director of the Arkansas Biosciences Institute. McGehee, a native Arkansan, joined the UAMS College of Medicine's Department of Pediatrics in 1993 and has received funding from the National Institutes of Health for his research on cellular differentiation and molecular mechanisms linking Type 2 diabetes and obesity. In 2012, UAMS honored Dr. McGehee by establishing The Robert E. McGehee, Jr., Ph.D. Distinguished Lectureship in Biomedical Research, supporting an annual lecture by a renowned scholar and leader in biomedical research and education.

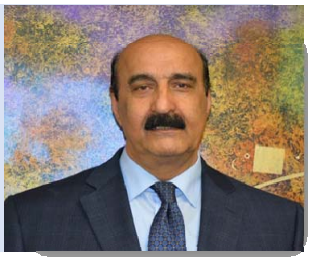
Ralph Davis, Ph.D.
Associate Vice Provost for Research and Innovation
And Professor in the Department of Geosciences
University of Arkansas
Fayetteville, AR

Ralph Davis is Associate Vice Provost for Research and Innovation and a Professor in the Department of Geosciences at the University of Arkansas. Ralph served as Chair of the Department of Geosciences for eight years (2008-2016), and prior to that he served seven years as Director of the Arkansas Water Resources Center (2001-2008), one of the 54 National Institutes for Water Resources. He earned degrees in geology and hydrogeology from the University of Nebraska. During his University of Arkansas tenure he has taught courses ranging from large lecture introductory geology to groundwater modeling at the graduate level. Ralph's research has focused on aspects of physical and chemical hydrogeology. He has been awarded 67 grants and contracts totaling over $4.3 million over his academic career. Ralph is a fellow in the Geological Society of America, and was awarded a Distinguished Service award by the Hydrogeology Division of GSA in 2005.

Abhijit Bhattacharyya, Ph.D.
Interim Vice Provost for Research and
Dean of Graduate School
University of Arkansas at Little Rock
Little Rock, AR

Dr. Abhijit Bhattacharyya began his academic career as an Assistant Professor in the Department of Mechanical Engineering, University of Alberta, Canada in 1997. Since January 2002, he has been a faculty member at the University of Arkansas at Little Rock, coordinator of the Applied Science graduate program (2006-2010), Associate Dean of the College of Engineering and Information Technology from 2011 to June 2016 (except during Spring 2015 when he served as Interim Dean). Since July 1, 2016, he has been serving as the Interim Vice Provost for Research and Dean of Graduate School. Since the beginning of his career, he has been involved in graduate education and has graduated seven PhD students and five MS students. His research is in the area of smart materials and thin films. He has been funded by the Department of Defense, Department of Energy, NASA and NSF. The American Society of Mechanical Engineers (ASME) recently recognized Dr. Bhattacharyya's significant contributions to the mechanical engineering community when he was elected Fellow of the American Society of Mechanical Engineers. He received his PhD in Mechanical and Aerospace Engineering from Rutgers University, New Jersey.

Mansour Mortazavi, Ph.D.
Vice Chancellor for Research, Innovation
and Economics Development
University of Arkansas at Pine Bluff
Pine Bluff, AR



In 1992, Dr. Mansour Mortazavi joined the University of Arkansas at Pine Bluff (UAPB). He began teaching at the university level as a graduate teaching/research assistant at the University of Arkansas from 1984 through September 2017. Dr. Mortazavi earned the rank of Full Professor in the Department of Chemistry and Physics and the Department of Mathematics and Computer Science.

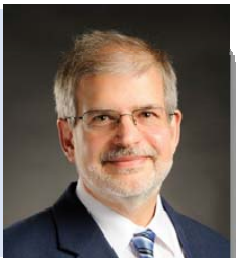
After joining UAPB, he began his research. Since 1995, he has continued to receive funding from federal and state agencies. As Principal Investigator, he has received awards from Air Force of Scientific Research; Army Research Laboratories; National Aeronautics and Space Administration (NASA); and National Science Foundation. Dr. Mortazavi has been in partnership with grants related to nanoscience, engineering and computer science disciplines.

He has publications featured in journals such as Science, Science News, Physical Review Letters, and Optics Letters. Dr. Mortazavi was involved in design and implementation of Spintronics research which had the world record for efficiency and consistency.

Dr. Mortazavi is affiliated with the University of Arkansas as a faculty member of Nanoscience and Engineering Institute and Micro-Electronics and Photonics. In September 2017, he was selected to serve as Vice Chancellor for Research, Innovation and Economic Development. Currently, he has initiated collaborations with the Pine Bluff Arsenal and the National Center for Toxicological Research in addition to partnerships with universities in the state of Arkansas.

Dr. Mortazavi is a member of several scientific societies and an honorary member of Sigma Pi Sigma and lifetime member of the Arkansas Academic Society.

Andrew Sustich, Ph.D.
Vice Chancellor for Research and
Executive Director of Biosciences Institute
Arkansas State University
Jonesboro, AR



Andrew Sustich is the Associate Vice Chancellor for Research and Executive Director of the Biosciences Institute at Arkansas State University, where he has been on the faculty since 1991. He holds an M.S. and Ph.D. in Physics as well as a B.S. in Nuclear Engineering, all from the University of Illinois.

Thomas Risch, Ph.D.
Professor of Animal Ecology
Arkansas State University
Jonesboro, AR

In 2001, Dr. Thomas Risch joined Arkansas State University (A-State) as a professor of Animal Ecology, after receiving his Ph.D. in Zoology from Auburn University. Dr. Risch has advanced through the promotion and tenure process at A-State and was promoted to the rank of Professor in 2011. He served as Director of the Ph.D. program in Environmental Sciences from 2009 through 2013 and as Chair of the Department of Biological Sciences from 2010 until 2018. Risch currently serves as the Interim Associate Vice Chancellor of Research & the Interim Executive Director of Arkansas Biosciences Institute at Arkansas State University. He holds the Judd Hill Endowed Chair of Environmental Biology.

Dr. Risch's research focuses on the ecology and conservation of mammals with a recent focus on bats in North America,but has published on a wide variety of taxa and topics including on human elephant conflict in Nepal. He has over 60 publications to date and maintains an active research lab. Risch's research has been funded by a variety of state and federal agencies including the U.S. Fish and Wildlife Service, U.S Forest Service, the U.S. Department of Energy, and the National Science Foundation.

AR-BIC 20119 Session Chairs Biographies

Argelia Lorence, Ph.D.
James and Wanda Lee Vaughn Endowed Professor
Director, A-State Phenomics
Co-Lead Wheat and Rice Center for Heat Resilience

The most significant contribution Dr. Lorence has made to plant sciences has been the discovery of a novel biosynthetic pathway for vitamin C that involves myo-inositol as a main precursor. Her laboratory uses Arabidopsis to better understand the role of various subcellular pools of vitamin C in plant physiology. Her ongoing research has potential applications for the development of crop plants with enhanced nutritional content, better growth, and improved tolerance to multiple environmental stresses. In addition to Arabidopsis, her current models of study include rice, soybean, and maize. Lorence directs the Plant Phenomics facility at A-State. She also co-leads the Wheat and Rice Center for Heat Resilience (WRCHR) funded by the NSF EPSCoR Track 2 program. This is a research consortium focused on finding novel genes involved in conferring rice and wheat tolerance to high night temperature, one of the key factors limiting the yields of the two most important crops in the world for food security. From 2014 to 2017 Lorence co-led the Plant Imaging Consortium (PIC), also funded by NSF EPSCoR. Since joining A-State in 2005, Dr. Lorence has secured over $18 million in grants. She is the President Elect of the Phytochemical Society of North America (PSNA).

Bryan J. Barnhouse , MPA
Vice President,
Arkansas Research Alliance
Conway, AR

Bryan comes to the Arkansas Research Alliance (ARA) as an experienced economic developer. In his previous six years with the Economic Development Alliance for Jefferson County (Arkansas), he directed business recruitment and retention, which included landing a $3.7 billion project that converts natural gas to clean diesel on more than 1,000 acres, and led programs that re-oriented the research and curriculum of higher education institutions to match corporate needs in the area. Prior to that, he managed the promotion of industrial recruitment from Asia and global trade and export development for the Arkansas Economic Development Commission.

Before relocating to Arkansas, he spent four years at the International City/County Management Association in Washington, D.C., coordinating federal business development activities and managing military and technology projects. His educational background includes a Master of Public Administration with an emphasis on Intergovernmental Management and a Bachelor of Arts in International Relations from the University of Southern California.

Bryan is committed to helping advance his adopted state of Arkansas through the creation of economic development opportunities. He is excited to be pursuing this work with ARA by connecting the state's rich scientific and research communities to each other and to the Arkansas business community. He helps professionalize the practice of economic development as a member of the board of directors for Arkansas Economic Developers. And, as a nine-year member and former Chair of the Little Rock Sister Cities Commission, he helps cultivate relationships and partnerships between Little Rock and cities around the world, particularly with Newcastle Upon Tyne, England, for which he serves as the Commission co-liaison.

Bryan and his wife, Jennifer, reside in the vibrant South Main area of downtown Little Rock and enjoy being part of its growth. They, along with their two rescue dogs, love to explore the outdoors of the Natural State.

Douglas D. Rhoads, Ph.D.
Director, interdisciplinary graduate program in Cell and Molecular Biology
University Professor of Biological Sciences
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Rhoads is director of the interdisciplinary graduate program in Cell and Molecular Biology, and University Professor in Biological Sciences, at the University of Arkansas. His research has focused on genomic analyses in a variety of species including human, chicken, tomato, bear, scorpion, yeast and bacteria. His primary research is in metabolic diseases affecting meat type chickens, and he is an affiliated faculty in the Center of Excellence for Poultry Science at the University of Arkansas. Currently funded projects are working on genomic mapping of genes affecting pulmonary hypertension, and etiology and epidemiology of bacterial chondronecrosis with osteomyelitis leading to lameness. Dr. Rhoads teaches a course in Genomics and Bioinformatics. Dr. Rhoads was a founding member of the Cell and Molecular Biology program (the largest Ph.D. program at the University of Arkansas), and has served as the director since 2006. Dr. Rhoads research has produced 58 journal articles, 11 industry technical reports, and 149 presentations.

Ohgew Kweon, Ph.D.
Staff Fellow, Division of Microbiology
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Ohgew Kweon is a research microbiologist of Division of Microbiology at FDA's National Center for Toxicological Research (NCTR/FDA). He has served science review/editorial boards and published about 40 papers and book chapters. He has a main research interest in bridging the gap between genome and phenome. Pleiotropy and epistasis of biological data are key factors for the successful genome-phenome mapping, essential for a better understanding of the microbial world. His research activities concern the development and application of systemic methodologies to address questions in the pleiotropic and epistatic complexity of biological data. Recently he and colleagues introduced a method called Network Based Functional Pan-Genomics (NBFPG) which systematically integrates the three different types of concepts: network, pan-genomics, and functional genomics. This approach allowed for phenotype-related functional pan-genomic comparison in the Mycobacterial Phenotype Network (MPN) at the genus level, which discovered pleiotropy- and epistasis-dependent evolutionary trajectories of the "PAH-degrading" phenotype in the genus *Mycobacterium*. Pleiotropy and epistasis are open concepts for all biology questions.

Shraddha Thakkar
Principle Investigator, Division of Bioinformatics and Biostatistics
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Thakkar works at FDA's National Center for Toxicological Research. Her research interests are in applying bioinformatics and cheminformatics for study of toxicity and drug development with specific interest in drug-induced liver injury. She has received multiple research and leadership awards regionally and nationally and with FDA. That includes Genentech Innovation in Biotechnology Award from American Association of Pharmaceutical Scientist (AAPS), Margret C. Etter Student lecturer award from American Crystallography Association, and Outstanding Service award from FDA. Dr. Thakkar has adjunct appointments at both University of Arkansas for Medical Sciences and University Arkansas at Little Rock (Assistant Professor). Furthermore, Dr. Thakkar was elected as Board member of the Mid-South Computational Biology and Bioinformatics Society (MCBIOS) in 2014 and served as President for the Society from 2016-2017. She is also the Chair of Pharmacogenomics Group (2018-19) and serves on the Awards Committee at American Association of Pharmaceutical Scientist (AAPS).

Steven Foley, Ph.D.
Deputy Director, Division of Microbiology
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Steven Foley is a Supervisory Research Microbiologist and Deputy Director of the Division of Microbiology at FDA's National Center for Toxicological Research (NCTR) in Jefferson, Arkansas. He earned my B.S. in Zoology and Ph.D. in Cellular and Molecular Biology/Infectious Diseases from North Dakota State University in Fargo. After his Ph.D. studies, he was a postdoctoral fellow with FDA's Center for Veterinary Medicine. Prior to joining NCTR, Dr. Foley was an Assistant Professor at the University of Central Arkansas and an Associate Research Scientist at the Marshfield Clinic Research Foundation in Wisconsin. Dr. Foley's research interests are largely in the fields of bacterial pathogenesis and antimicrobial resistance among bacterial foodborne pathogens and understanding the distribution of microbial populations in FDA-regulated products. Through these studies, Dr. Foley's research has utilized several bioinformatics and computation biology approaches to better understand the biological results obtained in the lab. Dr. Foley is also an Adjunct Professor in the Food Science Department and a member of the Cell and Molecular Biology graduate faculty at the University of Arkansas. He currently serves on several FDA-wide committees and is a member of the Interagency Risk Assessment Consortium and the Health and Environmental Sciences Institute's Microbiome Subcommittee.

Weida Tong, Ph.D.
Director, Division of Bioinformatics and Biostatistics
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Weida Tong is Director of Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR/FDA). He has served science advisory board for several multi-institutional projects in Europe and USA. He also holds adjunct appointment at several universities. In addition, he is the founder and board chairperson of newly established international MAQC Society. His division at FDA is to develop bioinformatic methodologies and standards to support FDA research and regulation and to advance regulatory science and personalized medicine. The most visible projects from his group are (1) conducting the Microarray and Sequencing Quality Control (MAQC/SEQC) consortium to develop standard analysis protocols and quality control metrics for emerging technologies to support regulatory science and precision medicine; (2) development of liver toxicity knowledge base (LTKB) for drug safety; (3) *in silico* drug repositioning for the enhanced treatment of rare diseases; and (4) development of various tools such as ArrayTrack™ suite to support FDA review and research on pharmacogenomics. In addition, his group also specializes in molecular modeling and QSARs with specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book chapters.

AR-BIC 2019 Speakers Biographies and Abstracts

Youngbeom Ahn, Ph.D.
Division of Microbiology
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Youngbeom Ahn is a research microbiologist in the Division of Microbiology at NCTR. His research focused on projects of FDA regulatory science including antimicrobial residues in foods and the safety of pharmaceutical products as related to microbial contamination. In cooperation with FDA's Center for Veterinary Medicine (CVM), he developed a project to assess the impact of antimicrobial residues on the human gastrointestinal tract microbiota. His research has provided information on methodological questions that have concerned national regulatory authorities in *in vitro* testing to determine if concentrations of veterinary antimicrobial agent residues entering the human colon remain microbiologically active. His research resulted in data of high significance that has assisted FDA and other national regulatory authorities in the International Cooperation on Harmonization of Technical Requirements for Registration of Veterinary Medical Products (VICH) Expert Working Group organized by CVM to evaluate the VICH GL36 guideline on human safety of veterinary antimicrobial drugs. He also initiated collaboration with FDA's Center for Drug Evaluation and Research (CDER) to develop a project which was funded by CDER to explore strategies for resuscitation and enrichment of *Burkholderia cepacia* complex strains in pharmaceutical products.

Recent Data on the Impact of Veterinary Drug Residues in Food Animal Derived Products on the Human Intestinal Microbiome.

The use of antimicrobials, such as tetracycline, in food-producing animals may result in antimicrobial drug residues (ADR) in edible tissues above permitted threshold levels from treated animals. ADR cause potential risk to consumers including the disruption of the intestinal microbiome and the emergence of antibiotic resistant bacteria. The Veterinary International Conference on Harmonization (VICH) document (VICH GL36(R)/FDACVM Guidance for Industry#159) provides guidance for assessing human safety of residues from veterinary antimicrobial drugs in edible foods of animal origin. International committees and national regulatory bodies have used a formula-based approach for determining microbiological acceptable daily intake (mADI) for antimicrobial drugs. The results of *in vitro* test application to determine A) no observable adverse effect levels (NOAELs), and B) the fraction of an oral dose of ADR available to intestinal microorganisms are presented in this study. We determined if tetracycline, at low residue concentrations, could impact the human intestinal microbiome structure and the tetracycline resistance-gene profile, following acute and subchronic exposure. The effects of 0.15, 1.5, 15, and 150 mg/ml of tetracycline (corresponding to 25, 250, 2500 and 25,000 µg/kg bw/day), after 24 h and 40 days of exposure, in 3% human fecal suspensions, collected from three individuals (A, B, and C) were investigated using *in vitro* batch cultures. Bacterial community analysis using rRNA-based pyrosequencing revealed that *Firmicutes* and *Bacteroidetes* were the predominant phyla in the three fecal samples. The ratio of phylotypes varied among individuals. The genus *Bacteroides* (of the

Bacteroidetes) was consistently increased at tetracycline concentrations of 0.15 µg/ml or above in individual A from 1.68 to 5.70% (24 h) and 4.82-8.22% (40 days), respectively, as well as in individual B from 5.13-13.50% (24 h) and 10.92-22.18% (40 days), respectively. Among the 23 tetracycline resistance genes (TRGs) screened, four *tet* genes (*tet*O, Q, W, and X) were major TRGs in the control and tetracycline-dosed fecal samples. A slight inconsistently increasing copy number of TRGs appeared to be related to tetracycline treatment, interindividual variability and duration of exposure. High-performance liquid chromatography and liquid chromatography mass spectrometry assays showed that 25% (w/v) diluted steam sterilized feces dosed with 0.15 and 1.5 µg/ml tetracycline had binding of 58.2 ± 10.8% and 56.9 ± 9.1%, respectively. Based on data from chemical and microbiological assay methods, the fraction of dose available to microorganisms was 0.418 and 0.431 of the 0.15 and 1.5 µg/ml tetracycline treatments, respectively. This study has helped assess the safety of veterinary antimicrobial-drug residues in food, their impact on the human intestinal microbiome, and the potential for antimicrobial resistance.

Chase Rainwater, Ph.D.
Associate Professor
Department of Industrial Engineering
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Chase Rainwater is an Associate Professor of Industrial at the University of Arkansas. He joined the faculty in August 2009. He currently serves as the Director of the J.B. Hunt Innovation Center of Excellence and Co-Director of the Arkansas Security Research and Education Institute. Dr. Rainwater's work has been supported through research funding from sources including the National Science Foundation, U.S. Department of Transportation, U.S. Department of Defense, U.S. Army Corps of Engineers and multiple industry collaborators. His research interests include supply chain optimization, large-scale algorithm design and information, transportation and food security. He has advised 5 Ph.D. dissertations along with multiple M.S. and undergraduate theses. He is an active member in the Institute for Operations Research and Management Sciences and Institute for Industrial and Systems Engineers. In the Northwest Arkansas community, Dr. Rainwater contributes frequently to STEM initiatives for K-12 students. Dr. Rainwater most recently was presented with the 2017-2018 Dean's Award of Excellence Collaborative Faculty Research Award for the College of Engineering. He received his Ph.D. in Industrial and Systems Engineering from University of Florida and BSIE degree from the University of Arkansas.

From Blockchain to the Internet of Things: Big Data and Its Role in Food Science:

The notion of a data revolution is being discussed across almost every industry across the country. With this come the questions of what is "big data," how do non-computer professionals benefit from it and what are the challenges and concerns to be considered? In this talk we will explore how all of these questions are relevant to the food science segment of our economy. We will discuss how the root of many learning-based data analysis advancements, predictive analytics, has broad applicability in the food science domain and applications that are already making use of such techniques. We will demystify the fundamental concepts that comprise machine learning. The differences between supervised and unsupervised learning, as well as classification will be illustrated. In addition, we will offer examples that illustrate the use of machine learning in industry via business-driven case studies. The roles of key commercial technologies, such as Blockchain, will be discussed in the context of data security and sharing amongst collaborative organization. We will also explore how growth of the so-called Internet of Things is feeding the "big data" craze while also posing numerous security challenges.

Dale Thompson, Ph.D., P.E.
Associate Professor
Department of Computer Science and Computer Engineering
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Dale Thompson is an Associate Professor with the Department of Computer Science and Computer Engineering (CSCE) at the University of Arkansas in Fayetteville, Arkansas. His research interests include computer networking, the Internet, cybersecurity, network security, and food defense. He teaches Computer Networks, Operating Systems, Wireless Systems Security, and Network Security. Currently, he leads the Food and Cyber Education (FACE) funded by NIFA/USDA to prepare graduate food science students to protect and defend food systems from cyberattacks. In addition, he has been leading the Training Arkansas Computing Teachers (TACT) project funded by NSF that is researching the best way to train high school teachers to teach computer science. He is a co-founder of the Arkansas Research and Education (ASCENT) Institute, whose primary mission includes security issues and technologies in cyber, transportation, critical infrastructure, and food systems. Past cybersecurity projects include IPv6 security, mobile anonymous communications, localization of radio frequency identification (RFID) tags, detecting counterfeit RFID tags based on signal fingerprinting, and developing modules for teaching RFID security.

Cybersecurity: Why Should I Care?

Personally identifiable information (PII) continues to be stolen from various data consolidators enabling identity theft scams. Ransomware and similar attacks have turned hacking into a profitable business by encrypting users' files and asking for payment. Malicious cryptominers are now infecting computing devices on the Internet to mine cryptocurrency for profit. Internet of Thing (IoT) devices are being deployed worldwide with security vulnerabilities and little economic incentives for patches and upgrades. Email phishing in which attackers attempt to obtain sensitive information by acting like a trustworthy entity continue to be common. Cybersecurity is the protection of Internet-connected computer systems from attacks on the hardware, software, and often data. Every organization connected to cyberspace must address cybersecurity challenges to successfully meet their mission and protect their business processes. To keep the nation secure, requires a cybersecurity workforce capable of designing, developing, implementing, and maintaining both defensive and offensive capabilities. This workforce includes not only technical staff, but managers that understand cybersecurity to implement the strategies and manage the risk in the organization. In this presentation, recent cybersecurity threats will be explained, initiatives to increase the number in the cybersecurity workforce will be presented, and good security practices to protect users' computer, network, and data will be suggested.

Daniel Berleant, Ph.D.
Professor, Department of Information Science
University of Arkansas at Little Rock

Dr. Daniel Berleant has served on the faculty at UA Little Rock since 2006. He is presently the Coordinator of the Master of Science in Information Science degree program, has served until recently as the coordinator of the Technology Innovation Graduate Certificate program, and was interim department chair in 2008. Prior to his current position at UA Little Rock he was an Associate Professor of Electrical and Computer Engineering at Iowa State University, and before that an Assistant and then Associate Professor in the Department of Computer Science and Computer Engineering at UA Fayetteville. A faculty member of the UA Little Rock / UA for Medical Sciences Joint Graduate Program in Bioinformatics, he was 2009-2010 President of the MidSouth Computational Biology and Bioinformatics Society (MCBIOS). His research activities have spanned bioinformatics, text mining, arithmetic on imprecise probability distributions, and technology foresight. Dr. Berleant has published over 100 articles and is the author of the book *The Human Race to the Future: What Could Happen — and What to Do* (4th ed. 2018).

From Better Corn to New Plant Paradigms

Increasing world population and climate change are driving unprecedented demands on limited land and water resources. Improving the yield and nutritional value of important crops such as corn (*Zea mays*) supports increased food quality and quantity using scarce land and water resources. The research goal of this project is to enable improving such crops by deepening our understanding of genetic and epigenetic mechanisms for protein accumulation in corn seed. We seek to understand the genetic factors that control protein accumulation during embryo maturation in corn using transgenic genotypes expressing a manganese peroxidase gene (MnP) driven by the maize globulin-1 promoter with the protein targeted to the apoplast as a window into gene regulation during embryo maturation in normal and directed corn genotypes. This will enable better understanding of cellular mechanisms of protein accumulation so that more efficient creation of high protein strains can be made. Understanding which genes/factors are critical and how they work together to influence the whole picture is our goal.

The project is directed by the Hood lab at Arkansas State University with the bioinformatics aspects of the project centered at the University of Arkansas at Little Rock. We plan to use combined bioinformatics and biotechnology techniques to seek an understanding of the multi-omic aspects of transgenic corn. Insights into the genomic regulatory mechanisms at play may further illuminate cellular and proteomic cascades, feedback loops, and related mechanisms for protein accumulation. Ultimately this is expected to improve nutritional quality of grain crops.

The specific aims of UA Little Rock's portion of the overall project are as follows. (1)*Transcriptome Variation Analysis.* Compare multi-step sequencing data analysis for different corn plant genotypes (i.e. control, low expressing, and high expressing). (2) *Promoter Effects on Proteomic Profile.* Determine if there are off-target effects on the proteomic profile related to the transgene compared to the native gene other than the expression of the target protein. (3) *Prediction of Protein Subcellular Localization.* Apply advanced machine learning techniques to perform sequence-based prediction of protein subcellular localization in differing genotypes. (4) *Lnc-RNA and small RNA involvement in Transcription Targets and Expression Profiles.* Discover if the interactions of Lnc-RNAs and other small RNAs with key regulatory proteins have predictable downstream effects on expression in the engineered corn plant.

The goals and plans of this project exemplify how desired plant properties will be achievable as biotechnology produces increasingly dramatic results. This talk will include comments on the future implications of this type of research.

Errol Strain, Ph.D.
Director, Biostatistics and Bioinformatics Staff/OAO
Center for Food Safety and Applied Nutrition
U.S. Food and Drug Administration (FDA)
College, Park, MD USA

Dr. Errol Strain is the Director of the Biostatistics and Bioinformatics Staff (BBS) in the Office of Analytics and Outreach (OAO) at the FDA Center for Food Safety and Applied Nutrition (CFSAN). He received his Bachelor of Science degree in Biochemistry from Purdue University in 1998 and his Ph.D. in Bioinformatics from North Carolina State University in 2006. The CFSAN BBS group provides comprehensive statistical and bioinformatics support to all internal and external regulatory and research-related CFSAN activities, including study design, analysis, and participation in the development and application of regulatory standards and compliance activities.

GenomeTrakr and GalaxyTrakr – Distributed Sequencing and Bioinformatics for Food

In 2019 whole-genome sequencing (WGS) will completely replace legacy technologies for detecting, tracking, and characterizing common bacterial pathogens in the food supply and for surveillance of foodborne illness in the United States. The influx of sequencing-based bio-surveillance data has led to an urgent need for tools not just to analyze sequence data, but to foster group collaboration and consistent interpretation of analysis results. In response, the Center for Food Safety and Applied Nutrition (CFSAN) at the Food and Drug Administration has deployed an instance of Galaxy called GalaxyTrakr; hosting a curated set of tools for microbial sequence analysis to empower collaboration and data sharing among food safety partners. Public health and agricultural laboratories, along with regulated industry, need to have an open, common platform to assess the quality of their WGS data and perform analyses to look for matches between isolates from different sources, and to look for markers associated with virulence, pathogenicity, and anti-microbial resistance. The approach deployed by CFSAN uses an elastic cloud in Amazon-Web Services to host a Galaxy server supporting laboratories and groups associated with GenomeTrakr. The tools and workflows developed for GalaxyTrakr can then be installed or imported into other Galaxy systems, allowing industry and third-party laboratories to analyze private data using the same tools as the FDA. Additionally, many tools developed for GalaxyTrakr are pathogen-agnostic, allowing public health partners to upload and analyze wide range of bacterial pathogens beyond *Salmonella,* STEC*, and Listeria*. This approach is already improving food safety while it hopes to allow for more rapid uptake of sequencing in other disease areas, thus benefitting all public health.

Eugene Young
Regional Director, Delta Regional Office
National Agricultural Statistics Service (NASS)
U.S. Department of Agriculture (USDA)
Arkansas, USA

Eugene Young is the Regional Director at USDA's National Agricultural Statistics Service (USDA/NASS) Delta Regional Office which services the states of Arkansas, Louisiana, and Mississippi. He has over 28 years of experience working for NASS. Mr. Young has previously worked in the Arkansas and North Carolina NASS field offices along with working in Washington, D.C. for 5 years. He is a graduate of the University of Arkansas and is a native of the state. His regional office's mission at USDA/NASS is to provide timely, accurate, and useful statistics in service to agriculture for the states of Arkansas, Louisiana, and Mississippi.

Agricultural Data – Supplying Information and Data Needs

Data needs in the agricultural sector are as diverse as the industry itself. As the statistical agency for the U.S. Department of Agriculture for more than 150 years, NASS are the official source of primary, comprehensive, current information on the farms, ranches, and people who provide food, feed, and fiber to our nation and the world. Data and measurements on all agricultural activities are made down to the local county level for all who are involved in production agriculture. NASS collects, assembles, processes, and disseminates data on all aspects of U.S. agriculture based on survey, satellite, and administrative information on American agriculture. This process is completed by conducting hundreds of national weekly, monthly, quarterly, and annual surveys each year, along with many more at regional, state, and local levels. Along with these sampled surveys, NASS is also charged with conducting a detailed Census of Agriculture every five years for each farm, ranch, and agricultural producer in the nation. Data from the Census of Agriculture paints a complete picture of agricultural activities along with a demographical makeup of the farming sector in America. On an annual basis, NASS will publish more than 400 national reports and hundreds of thousands of data items to help others make farm-level, business, and policy decisions. NASS also partners with state agriculture departments, universities, other federal agencies, and others to conduct additional surveys to meet partners' specific needs. Nass conducts ongoing statistical research on survey design, sampling and other topics to advance the accuracy of statistical science. All data published are available free of charge and all data reports are released at dates and times published a year in advance. NASS is recognized internationally for our agricultural expertise and resources, the quality of our data, and our contributions to statistical science. We work with the U.S. Agency for International Development and other federal agencies to provide technical assistance and training to help developing and transitioning countries improve and expand their capacity to produce agriculture statistics.

Fiona L. Goggin, Ph.D.
Professor, Department of Entomology
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Fiona Goggin has been a faculty member at UAF since 2001, and her primary area of expertise is pest resistance in plants. Recent research areas include the influence of fatty acid metabolism, reactive oxygen species, and peptide signaling in plants on defenses against aphids and root-parasitic nematodes. She also teaches a graduate-level class in insect physiology, and was the recipient of the 2014 Southeastern Branch Entomological Society of America Recognition Award in Insect Physiology, Biochemistry and Toxicology. Dr. Goggin has served as a panel manager for USDA-NIFA, an editorial board member for the Journal of Chemical Ecology, and a guest editor for *Current Opinion in Insect Science* and *Agronomy*.

Applications of Imaging Technologies and Materials Science to Plant Phenomics

Together, all the facets of an organism's appearance and behavior equal its phenome---the sum total of all of its many phenotypes, such as size, shape, growth rate, disease resistance, and so forth. An organism's genome dictates the range of potential phenotypes it may develop, depending upon its environment. Due to relatively recent advances in DNA sequencing technologies, full genome sequences are now available for a wide range of living things, including many crop plants. Next, a major thrust in 21st century bioscience is to align these genomes with phenomes in order to determine which genes are responsible for which phenotypes, and to predict potential phenotypes based on genomic information. In fact, decoding the relationships between genomes and phenomes is one of the five Grand Challenge areas in Biology identified by the National Research Council and the National Science Foundation (NSF). As technologies for characterizing whole genomes have advanced, methods for characterizing phenomes have lagged behind, and phenotyping has become a major, rate-limiting step in the process of determining the functions of each component of a genome. In the field of plant science, the Plant Imaging Consortium (PIC) and Engineering Research Center for Materials for Agriculture Resource Imaging Analytics at High Resolution (MARIAH) have formed to help address this problem. With funding from NSF**, PIC** brought together experts in plant biology, radiochemistry, phenomics, imaging, and computational biology to apply high-throughput phenotyping and molecular imaging techniques to the study of plant stress biology. This consortium worked to make infrastructure for high-throughput plant phenotyping (HTPP) and molecular imaging (MI) accessible to plant biologists throughout Arkansas and Missouri, develop new protocols and analysis tools to apply HTP and MI to plant biology, and promote community standards for the design, analysis, and reporting of phenomics data. As an extension and expansion of this effort, the PIC community is currently partnering with experts in nanotechnology to form an engineering center dedicated to sensor development for plant phenotyping: the Engineering Research Center for Materials for Agriculture Resource Imaging Analytics at High Resolution (MARIAH). Technologies developed by MARIAH would enable high-throughput, non-invasive plant phenotyping, as well as improved monitoring of the environmental conditions that interact with genotypes to shape phenotypes.

Giovanni Petris, Ph.D.
Professor and Director of Statistics
Department of Mathematical Sciences
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Giovanni Petris is Professor and Director of Statistics in the Department of Mathematical Sciences at the University of Arkansas. His main research interests include computationally intensive methods for statistical inference, Bayesian statistics, and time series analysis. A recognized expert in the statistical software R, he is the author of the popular R package dlm for the analysis of time series using state space models, and of the book "Dynamic linear models with R." Dr. Petris has collaborated on research projects with the US Forest Service and the US Energy Information Administration. He has acted as a statistical consultant for many major companies. When he is not doing statistics, he can be spotted running on the beautiful Arkansas trails or playing chess.

Practical hierarchical Bayesian modeling using R

The workshop provides an introduction to the basic tenets of Bayesian inference, including construction of priors, summarization of the posterior, prediction and model checking. There will be a focus on regression models for continuous and categorical response data, including multilevel models where Bayesian hierarchical models provide a convenient alternative to classical random effects models. The use of R in Bayesian computations will be described. In particular, we will discuss the application of Markov chain Monte Carlo algorithms to Bayesian inference and diagnostic tools to assess the convergence of the MCMC sampler. For the more advanced examples we will use Stan and the accompanying rstan package to fit Bayesian models via MCMC algorithms. It is assumed that the participant has a basic familiarity with the R software environment.

Ioannis Tzanetakis, Ph.D.
Professor, Department of Entomology and Plant Pathology
Division of Agriculture, University of Arkansas

Dr. Tzanetakis is a Professor of Plant Virology at the University of Arkansas and the Director of the Arkansas Clean Plant Center for Berries. He has a BS in Soil Science and Agricultural Chemistry and a PhD in Molecular and Cellular Biology working on berry viruses. He has worked on translational mechanisms of RNA plant viruses and bioinformatics analysis of coupled translation in eukaryotes. Currently his research focuses on the epidemiology and detection of plant viruses and particularly on the development of vertical pipelines starting from wet lab protocols to custom-designed bioinformatics tools for detection and discovery of pathogens using high throughput sequencing. Dr. Tzanetakis is a member of the editorial boards of several plant pathology and virology-focused journals and holds leadership roles in several national and international groups including the American Phytopathological Society, the National Clean Plant Network, the International Council for the Study of Virus and other Graft Transmissible Diseases of Fruit Crops and the International Committee on Taxonomy of Viruses.

How bioinformatics impact plant movement across border lines

High throughput sequencing (HTS) has revolutionized detection and discovery of plant viruses and viroids, with more than 200 new species identified in the last few years; affecting the movement of propagation material across state and country lines. There are major bottlenecks in the process; the extraction methodologies and template preparations (ribodepleted RNA, dsRNA, small RNAs, or virus particles) and the virus detection and discovery bioinformatics pipelines which may not have the capacity to identify new agents based on structural motifs (e.g. viroids) or conserved protein domains. Yet, possibly the most important factor in the process is human interpretation of the results. An experienced plant virologist is always needed to interpret the output of the pipelines and distinguish between noise and what needs to be further investigated as a potentially new pathogen. Even with all the bottlenecks, HTS has been successfully tested against conventional tools such as bioindicators, ELISA and PCR and has become evident that it will be the method of choice not only for plant virology research purposes but also for plant certification programs across the globe. This presentation will focus on the attributes of the different detection technologies and the translational application of HTS on certification rules and regulations in the United States and across the globe.

Jackson Cothren, Ph.D.
Director, Center for Advanced Spatial Technologies
Director, Arkansas High Performance Computing Center
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Jackson Cothren is Director of the Center for Advanced Spatial Technologies (CAST) and the Arkansas High Performance Computing Center (AHPCC) at the University of Arkansas. He is also appointed as a Professor of Geosciences in the Department of Geosciences at the University of Arkansas. CAST is a multi-disciplinary center for spatial research and technology housed within the J. William Fulbright College of Arts and Sciences at the University of Arkansas. CAST focuses on applications of geospatial technologies in research, teaching, and service. With expertise in the measurement and analysis of spatially referenced, multi-scalar features and processes, CAST collaborates with researchers at the various universities and various research organizations to extend and enhance research in various social and physical science domains. AHPCC provides expertise, high performance computing hardware, storage, support services, and training to enable computationally-intensive and data-intensive research. The AHPCC is available to faculty, staff and students at all of the Arkansas public universities, and to their collaborators inside and outside of the state. Most AHPCC services are provided free of charge to eligible researchers. Dr. Cothren's primary research interests include various aspects of digital photogrammetry and computer vision including geometric sensor modeling, point-cloud extraction from imagery, feature extraction and matching for orientation, integration of LIDAR point-clouds with imagery, error propagation and reliability analysis in adjustment models and implementation.

Tragedy of the Commons? Lessons in Sharing Distributed Scientific Instruments

As the number and variety of scientific instruments across university campuses, state and federal laboratories, and private interests grows, the ability to remotely oversee or monitor experiments and analyze collected data during and after the experiment is a potential limiting factor in their widespread, shared use. The NSF EPSCoR-funded Plant Imaging Consortium encountered this potential problem with the deployment of the Lemnatec Scanalyzer at Arkansas State University in Jonesboro. The Scanalyzer is a lab-based, high-throughput plant phenotyping instrument that can provide high spatial- and temporal-resolution multi-spectral images of differentially treated plants during the growth phase. Data from these experiments is collected daily or weekly over several months and often early problems with an assay can go undetected if the researcher does not carefully monitor the experiment. This is generally not a problem when the researcher and instrument are collocated. However, when a researcher is five hours away in Fayetteville, it is cost prohibitive to travel to the instrument throughout the experiment. Furthermore, the instrument control software is often built to run on a single workstation directly connected to the instrument and collected data is not automatically nor easily shared across a network. The multi-institution bioinformatics core of the Plant Imaging Consortium worked together to develop a

web-based interface for the Scanalyzer that 1) enforced MIAPPE meta-data standards about the experiment, 2) collected and reformatted meta-data associated with the collected images, and 3) immediately after collection moved local meta-data and data to a storage array shared across campuses. The combination of these capabilities makes is easier for remote researchers to monitor experiments and make modifications during the long growing period, review images immediately after collection, and use meta-data to more effectively organize further analysis. We will discuss the challenges associated with the effort and discuss a state-wide infrastructure to make this kind of scientific collaboration more effective.

James Reecy, Ph.D.
Associate Vice President for Research
Iowa State University
Iowa, USA

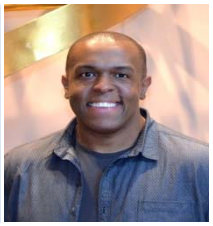
Dr. James Reecy currently serves as an Associate Vice President for Research where he oversees the Office of Sponsored Programs Administration and internal funding programs and fosters the development of interdisciplinary teams. He joined the faculty of Iowa State University in February of 1999 and is currently a Professor in the Department of Animal Science. He received a B.S. degree from South Dakota State University; a M.S. degree from the University of Missouri, Columbia; and a Ph.D. degree from Purdue University. He served as the Director of the Office of Biotechnology, which administered 10 fee-for-service core facilities for 10 years. Dr. Reecy currently is the NRSP-8 database coordinator, where he leads national efforts to improve the computational resources available for genomics research on livestock species. In addition, he is currently serving as a 2018 Association of Public and Land Grant Universities (APLU) Council on Research Fellow. During his career, Dr. Reecy has worked on problems in ruminant nutrition, skeletal muscle growth and development, embryonic heart development, beef and mouse molecular and quantitative genetics, and livestock bioinformatics. His lab has worked on beef cattle molecular genetics with a focus on improving the nutrient content of beef and health of cattle, as well as, the development of database resourced to facilitate genomics research.

Animal QTLdb and CorrDB: Helping to close the genotype to phenotype gap

Thanks to technological advances, animal geneticists have a massive tool chest with which to study the inheritance of traits in livestock in order to improve production. The ability to identify genetic markers led to the establishment of linkage maps, which enabled the study of quantitative trait locus (QTL) association of traits with genome segments. The identification of single nucleotide polymorphisms (SNPs) gave rise to high-throughput genotyping with chips and, in turn genome-wide association study (GWAS), to illustrate the genotype-phenotype landscape in greater detail. Functional annotation of livestock genomes will likely accelerate research findings. All of these processes have generated unprecedented amounts of information for geneticists to digest and translate into understanding of the genetic architecture of animal traits. The Animal QTLdb (https://www.animalgenome.org/QTLdb) and Animal CorrDB (https://www.animalgenome.org/CorrDB) were created to help meet these challenges. To date, the Animal QTLdb has accumulated over 158,000 QTL/association data on nearly 2,000 animal traits in 6 species (cattle, chicken, horse, sheep, pig, and rainbow trout). The CorrDB contains 10,482 correlations on 376 traits and 1,984 heritability estimates on 621 traits in 4 animal species (cattle, chicken, pig, and sheep). We have used ontologies (Vertebrate Trait Ontology https://www.animalgenome.org/bioinfo/projects/vt/ and Livestock Product Trait Ontology https://www.animalgenome.org/bioinfo/projects/lpt/) to provide a convenient and consistent way by

which data can be classified, annotated, and compared among livestock species. We have also developed tools to help better understand the potential relationship between traits. For example, Gene Ontology enrichment analysis of large gene expression experiments has been recognized as an effective method for investigators to better understand the biological processes most pertinent to their studies. At an abstract level, a gene represents a region of the genome as does a QTL. Similarly, gene ontology terms have been associated with genes as phenotypes are with QTL. Therefore, we have developed tool to evaluate if regions of the genome are enriched for traits. With ongoing advancements in animal genome research, it is evident that continued developments and additional expansion of both databases are urgently needed to facilitate future research. Our *long-range goal* is to develop integrated resources that leverage prior investments in cyberinfrastructure to help researchers maximize the utility of genotype-to-phenotype data to ultimately address issues of importance to the livestock industry.

Jason Williams
Assistant Director, External Collaborations
DNA Learning Center
Cold Spring Harbor Laboratory
New York, USA



Jason Williams is Assistant Director, External Collaborations of Cold Spring Harbor Laboratory's DNA Learning Center, and is the Education, Outreach, and Training lead for CyVerse (A U.S. national life science cyberinfrastructure funded by NSF). Jason organizes, instructs, and speaks at a variety of bioinformatics-related workshops, conferences, and meetings annually. He also serves in an advisory capacity on a variety of bioinformatics and open science projects including his service as Chair of the International Science Advisory Board for EMBL-Australian Bioinformatics Resource, and service on the Science and Industry Advisory Board of ELIXIR UK, External Panel of Consultants to the National Institutes of Health (NIH) Data Commons Initiative, and the External Expert Panel of NIH's National Heart Lung and Blood Institute's Data STAGE (Storage, Toolspace, Access and analytics for biG data Empowerment). He is an active Software and Data Carpentry instructor, and a former Chair of the Software Carpentry foundation (an international organization of researchers that promote training and education in software development, scientific data management, and open science). Jason is also an instructor at the Yeshiva University High School for Girls.

CyVerse Cyberinfrastructure for Research and Education in Genomics and Metagenomics

CyVerse (http://www.cyverse.org/ - formerly *iPlant Collaborative)* is a project funded by the National Science Foundation to develop computational resources (software, data storage, high-performance computing, and people) for life science. CyVerse's web-based platforms provide access to 1) a personal data store and data commons, 2) cloud and HPC computing, and 3) graphical interface to bioinformatics applications, including a visual and interactive computing environment for Jupyter Labs and RStudio applications. In this workshop we will introduce powerful tools that support the data and computational needs of big data biology. In particular, we will demonstrate how to upload, share, and analyze data using RNA-Seq as an example workflow. Additionally, we will introduce DNA Subway, a classroom bioinformatics pipeline that allows students to predict and annotate genes in up to 150kb of DNA (*Red* Line), identify homologs in sequenced genomes (*Yellow* Line), identify species using DNA barcodes and phylogenetic trees (*Blue* Line), and examine RNA-Seq datasets for differential transcript abundance (*Green* Line). The new *Purple* Line is a graphical, user-friendly implementation of the QIIME2 workflow. This workflow will can be used to examine how 16s sequencing reveals microbial diversity in the context of student-driven research projects. All tools and resources are funded by NSF and freely available.

Jeremy Edwards, Ph.D.,
Research Plant Molecular Geneticist
ARS Dale Bumpers National Rice Research Center,
USDA,
Stuttgart, AR

Dr. Jeremy D. Edwards is a Research Plant Molecular Geneticist with USDA's Agricultural Research Service at the Dale Bumpers National Rice Research Center in Stuttgart, AR. His research involves the use of genomics and bioinformatics and big data to facilitate gene discovery and to develop technology for accelerated rice breeding. He has a Ph.D. in Plant Breeding and Genetics from Cornell University and a B.S. in Horticultural Science from the University of Florida.

Ricebase: A web Resource for Rice Breeding and Genetic Discovery

Ricebase (http://ricebase.org) is an integrative genomic database for rice (Oryza sativa) with an emphasis on combining datasets in a way that maintains the key links between past and current genetic studies. Ricebase includes DNA sequence data, gene annotations, nucleotide variation data and molecular marker fragment size data. Rice research has benefited from early adoption and extensive use of simple sequence repeat (SSR) markers; however, the majority of rice SSR markers were developed prior to the latest rice pseudomolecule assembly. Interpretation of new research using SNPs in the context of literature citing SSRs requires a common coordinate system. A new pipeline, using a stepwise relaxation of stringency, was used to map SSR primers onto the latest rice pseudomolecule assembly. The SSR markers and experimentally assayed amplicon sizes are presented in a relational database with a web-based front end, and are available as a track loaded in a genome browser with links connecting the browser and database. The combined capabilities of Ricebase link genetic markers, genome context, allele states across rice germplasm and potentially user curated phenotypic interpretations as a community resource for genetic discovery and breeding in rice.

Joshua Xu, Ph.D.
Branch Chief, Research-to-Review and Return (R2R)
Division of Bioinformatics and Biostatistics
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)

Dr. Joshua Xu is currently the Branch Chief for Research-to-Review (R2R) Branch at the Division of Bioinformatics and Biostatistics of NCTR. He has about twenty years developing bioinformatics software and systems and conducting bioinformatics research. He specializes in software design and implementation, big data analytics, and machine learning. At NCTR, he has led several system development projects including SNPTrack, which is an integrated solution for managing, analyzing, and interpreting genetic association study data. His recent endeavor has been with the Sequencing Quality Control project phase II (SEQC2), a large and international collaborative consortium led by the FDA to evaluate the technical reliabilities and scientific applications of the next generation sequencing (NGS) technologies. As the principle investigator (PI), he is leading a large Working Group as part of the SEQC2 Consortium to assess the reproducibility and detection sensitivity of onco-panel sequencing including liquid biopsy. The Working Group consists of over 200 participants from academia, government agencies, and industry including 8 companies providing onco-panels and 30 testing laboratories. The scope and complexity of this comprehensive study is unprecedented. The project aims to provide recommendation in support for FDA's mission in regulatory oversight of NGS diagnostic tests. He is also the principle investigator for two other research protocols at NCTR.

Developing an Intelligent System for Species Identification of Food Contaminating Beetle

Insect pests are often associated with food contamination and public health risks. Accurate and timely species-specific identification of pests is a key step to scale impacts, trace back the contamination process and promptly set intervention measures, which usually have serious economic impact. The current procedure involves visual inspection by human analysts of pest fragments recovered from food samples, a time-consuming and error-prone process. In this collaborative project with FDA food analysts at Arkansas Laboratories, we are developing an automated and intelligent system for species identification of such food contaminating beetles through image analysis and machine learning. 1) We demonstrate such feasibility with both artificial neural network (ANN) and support vector machine (SVM) based methods to identify food contaminating beetle species by using extracted features on their elytra (hardened fore wings). Both methods show good accuracies between 80% and 85% when horizontally cropped sub-images of whole elytra were used for the analysis. With unknown orientations of the fragments recovered from food screening, we observe that the accuracy dropped by ~10%, compared to the horizontally cropped sub-images. 2) We apply deep neural architectures to the same dataset of 15 storage product beetle species, obtaining an overall accuracy of 84% in cross validation.

Notably, the classification performance is obtained without the need of designing and selecting domain specific image features and random image orientation only lower the accuracy by ~5%. 3) Investigations reveal that the misidentification is more prone for images that lack the visual clarity due to the reflective glares of the elytra surface. Thus, we amend the optical and imaging settings by incorporating polarizer filters that reduce the blurring from glares. Further adjustments to light sources and their orientations allow us to capture images with much increased consistence and enhanced clarity. An imaging manual has been developed. 4) We are refining a prototype image analysis application with a graphical user interface to integrate all image analysis and review steps. This application will be tested by food analysts. In summary, our work addressed both imaging and image processing methods, necessary for accurate identification of food contaminating beetle species.

Lawrence J. Lesko, B.S., Ph.D., F.C.P.
Clinical Professor Emeritus and Founding Director
Center for Pharmacometrics and Systems Pharmacology
University of Florida College of Pharmacy
Lake Nona (Orlando), FL, USA

Dr. Lawrence J. Lesko is Clinical Professor Emeritus and Founding Director of the Center for Pharmacometrics and Systems Pharmacology at the University of Florida College of Pharmacy in Lake Nona (Orlando) Florida. He started the Center in 2011 and served as Director until April 2016. Dr. Lesko was appointed as the first Director of the Office of Clinical Pharmacology in the Center for Drug Evaluation and Research at the Food and Drug Administration in 1995 and served as Director until 2011. At the FDA, Dr. Lesko was a leader in personalized medicine, being responsible for the update of approved drug labels with genetic information for individualized drug selection and precision dosing. Dr. Lesko led a program at FDA that focused on mechanistic approaches to understanding drug safety in patients including the prediction of off target adverse effects using bioinformatics. He was also a leader in therapeutic drug monitoring services at the University of Massachusetts Medical Center and University of Maryland. Dr. Lesko has published more than 200 peer-reviewed scholarly manuscripts and is a frequent invited speaker at national and international meetings of professional organizations. From

Real World Evidence Drives Innovation in Healthcare

2006 to 2008, Dr. Lesko was President of the American College of Clinical Pharmacology.

This presentation will explore contemporary examples of real world data (RWD) and unique applications of real world evidence (RWE) to advance healthcare knowledge and decision-making. Cliches become clichés because they are often true. Gary King, Harvard University, makes the point: "Big data is not about the data – the real value is in the analytics". We have reams of healthcare data to prove the point that hiding within these data heaps is knowledge that can change the life of an individual patient. Real world data (RWD) fits into a broader trend of innovation in healthcare, which is individualized and tailored to medical needs and delivery of healthcare to patients. At the heart of current innovations – "wearables"— are three things: (1) a device, (2) a sensor, and (3) an algorithm. First generation fitness trackers (Fitbit®) and Apple Watch apps (KardiaBand®) keep tabs on step counts, heart rate, EKG and sleep rhythms. A new generation of devices that gather RWD and aim to assess a person's underlying biology are rapidly becoming reality. While everyone talks about RWD ("THE WHAT"), what makes innovation come to life is RWE ("THE HOW"). Today, wearable devices can take real-time snapshots of a person's EKG recording (RWD) and detect changes that predispose to heart attacks (RWE), tumor biopsies can reveal genetic changes (RWD) that determine whether a patient will benefit from chemotherapy (RWE), and cameras that can take high-resolution images of the retina (RWD) and detect early signs of glaucoma and macular degeneration (RWE). Someday soon, soft wearable patches will be used from "cradle to grave" to analyze electrolytes and glucose in sweat to warn about dehydration, screen for cystic fibrosis and self-monitor diabetes. Meanwhile, teams in the pharmaceutical industry

and in regulatory agencies such as FDA have laid out plans to use RWD to address the efficacy gap between randomized controlled trials and routine clinical practice, to inform important healthcare decisions regarding post-approval monitoring of adverse drug events and risk management, and to provide quantitative metrics to compare competing or alternative treatment approaches.

Noah Fahlgren, Ph.D.
Director, Bioinformatics Core Facility
Donald Danforth Plant Science Center
Missouri, USA

Dr. Noah Fahlgren is the Director of the Bioinformatics Core Facility at the Donald Danforth Plant Science Center. The Bioinformatics group uses and develops computational approaches and infrastructure that leverage large datasets to address biological problems. We emphasize the development of modular, reusable, and open-source tools through collaborator- and community-driven efforts. Our aim is to apply these tools to high-throughput genotyping and phenotyping data to identify the genetic basis of traits in research model plants and biofuel and food security crops.

The ability to rapidly and non-destructively measure plant physical and physiological features is a key bottleneck in plant research and breeding. Imaging coupled with computer vision algorithms and statistical analysis are a set of technologies that have the potential to address the plant phenotyping bottleneck, but they introduce their own computing, interpretation, and data management challenges that our group develops tools to address so that these technologies can be utilized more broadly by the scientific community. Plant Computer Vision (PlantCV) is our primary platform for developing a plant phenotyping toolbox. Through PlantCV we are deploying computer vision, machine learning, and other data science algorithms to extract biologically relevant data from image and sensor datasets.

Machine Learning Methods in PlantCV for Leaf Tracking and More

Machine learning methods have emerged as a set of powerful tools for extracting information from large datasets. While supervised methods require human-curated training data that can be challenging to collect, in the area of plant phenotyping they provide an opportunity to connect the domain expertise of plant biologists with flexible learning algorithms developed by computer scientists. For plant image analysis problems, deep learning approaches such as deep neural networks are particularly attractive solutions to extracting plant phenotype information that would otherwise be difficult to extract using traditional computer vision techniques. Many well-evaluated architectures for deep neural networks exist and can be applied to plant phenotyping problems through complete retraining or updated training. The choice of architecture depends on the type of output information that is desired. Output label types can include classification (what category does the image belong to), semantic segmentation (what pixels belong to each category), object detection (what are the bounding regions of each object of each category), and instance segmentation (what pixels belong to each object of each category). Here we demonstrate the use of an instance segmentation-class method, trained with publicly available data, for identifying and labeling individual leaves of Arabidopsis thaliana. The trained model performs well on a validation set drawn from the publicly available labeled data but did not immediately perform well on data from other sources that were not part of the public set. To address this we collected new training

data using a deep learning-augmented labeling technique and updated the leaf segmentation model to achieve better performance on the additional datasets. By using the segmentation method on time-series image datasets we are able to track the size of individual leaves over time using PlantCV (http://plantcv.danforthcenter.org).

Samantha Robinson, Ph.D.
Clinical Assistant Professor
Department of Mathematical Sciences
University of Arkansas
Fayetteville, Arkansas, USA

Dr. Samantha Robinson is a Clinical Assistant Professor in the Department of Mathematical Sciences at the University of Arkansas. She also serves as the course coordinator for all introductory statistics courses (2000 level) at the University of Arkansas, including both Biostatistics and Principles of Statistics courses. In addition, she serves as a Senior Applications Systems Analyst consulting with the Department of Family and Preventative Medicine (DFPM) Community Research group at the University of Arkansas for Medical Sciences (UAMS). Prior to her ongoing collaboration with the Community Research group at UAMS, she served as a statistical consultant for medical professionals located in both health care facilities and corporations such as Washington Regional Medical Center, the Cerner Corporation, and BioMérieux. Her primary research interests include psychometric modeling (especially in the context of large scale assessments), local spatial analysis, and statistics education.

Introduction to R

R is an open-source programming language for statistics and data science that has rapidly expanded, gaining traction and becoming a dominant fixture in both academia and industry. This introductory workshop is intended for those that have not used R for statistical analysis before but are interested in doing so.

Assuming no background knowledge of R, the workshop will cover the following:

- Installation of R and/or RStudio
- Script creation
- Data import and data structure basics
- Calculation of summary statistics
- Creation of visuals such as histograms, boxplots, and mean plots with error bars
- Linear Regression (Simple Linear and Multiple Linear) with accompanying visualizations
- ANOVA (One Way and Two Way Between Subjects) with accompanying visualizations *as time allows*
- Additional/Further Resources

With the knowledge gained in this workshop, the new user of R will have the requisite knowledge to conduct basic data analysis and will be ready to begin independently exploring all that R has to offer!

Steven C. Ricke
Wray Endowed Professor
Director - Center for Food Safety
Department of Food Science
University of Arkansas
Fayetteville, AR

Dr. Steven C. Ricke is a Professor with the Department of Food Science at the University of Arkansas in Fayetteville, Arkansas. He is also the Director of the Center for Food Safety and is the Donald (Buddy) Wray Endowed Chair in Food Safety. He is a co-founder of the Arkansas Research and Education (ASCENT) Institute, whose primary mission includes security issues and technologies in cyber, transportation, critical infrastructure, and food systems. Dr Ricke's research involves using a variety of molecular and genomic techniques with a primary focus on the understanding of the mechanisms of foodborne bacterial pathogen (*Salmonella*, *Listeria*, and *Campylobacter*) contamination at all phases of food production to develop a more integrated control effort. He has more recently initiated research on microbiome sequencing analyses of food production systems. Specific projects include assessment of poultry gastrointestinal tract microbial community responses to feed additives such as organic acids and prebiotics and application of microbiome mapping of microbial populations during poultry processing.

Food Production and Microbiome Applications

Food production involves multiple steps to generate the final food product for retail distribution. For example, during poultry processing, poultry carcasses are exposed to variations in moisture levels and application of antimicrobials along with other environmental conditions. Consequently, bacterial population composition can shift and this may be important for evaluating effectiveness of control measures designed to reduce bacterial contamination and improve shelf life of the final product. Monitoring poultry carcass bacterial loads has traditionally been done by conventional microbial plating and enumerating the colonies on these plates. Next generation sequencing of the microbiome has led to in-depth microbial taxa profiling and diversity comparisons among different microbial communities from different sources. Microbiome sequencing has been applied to poultry processing and the results have revealed major shifts in bacterial taxa during the processing steps with different groups becoming predominant depending on the specific processing step. As more becomes known about these microbial populations based on microbiome sequencing, opportunities to identify more representative indicator organisms and more precisely track microbial shifts in populations related to a specific processing step are becoming possible. As more microbiome data is generated and analyzed the concept of microbiome mapping may become a practical tool for routine monitoring of processing performance and effectiveness of control measures.

Steven Foley, Ph.D.
Deputy Director, Division of Microbiology
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Steven Foley is a Supervisory Research Microbiologist and Deputy Director of the Division of Microbiology at FDA's National Center for Toxicological Research (NCTR) in Jefferson, Arkansas. He earned my B.S. in Zoology and Ph.D. in Cellular and Molecular Biology/Infectious Diseases from North Dakota State University in Fargo. After his Ph.D. studies, he was a postdoctoral fellow with FDA's Center for Veterinary Medicine. Prior to joining NCTR, Dr. Foley was an Assistant Professor at the University of Central Arkansas and an Associate Research Scientist at the Marshfield Clinic Research Foundation in Wisconsin. Dr. Foley's research interests are largely in the fields of bacterial pathogenesis and antimicrobial resistance among bacterial foodborne pathogens and understanding the distribution of microbial populations in FDA-regulated products. Through these studies, Dr. Foley's research has utilized several bioinformatics and computation biology approaches to better understand the biological results obtained in the lab. Dr. Foley is also an Adjunct Professor in the Food Science Department and a member of the Cell and Molecular Biology graduate faculty at the University of Arkansas. He currently serves on several FDA-wide committees and is a member of the Interagency Risk Assessment Consortium and the Health and Environmental Sciences Institute's Microbiome Subcommittee.

Antimicrobial Resistance Dissemination Among Enteric Bacteria

Members of the family *Enterobacteriaceae* rank among the most common causes of bacterial foodborne illnesses and other infectious diseases in the United States. Among these enteric bacteria, antimicrobial resistance has been identified as a major public health threat worldwide. Many of the tools used to study antimicrobial resistance in enteric bacteria rely on the generation of large biological datasets and require effective bioinformatics tools to analyze the data. A key worry related to antimicrobial resistance is the potential for horizontal transfer of antimicrobial resistance factors among bacteria. A specific area of concern is whether enteric bacteria within the gastrointestinal tract can serve as a reservoir for genetic elements that encode antimicrobial resistance. Many antimicrobial resistance determinants are located on plasmids and other mobile genetic elements, such as integrons and transposons, that can be transferred among bacteria. Plasmids can also carry genes that contribute to increased virulence as well, thus it is important to better understand their roles in virulence and antimicrobial resistance in enteric organisms. We have extensively studied plasmids that are defined as being members incompatibility groups (Inc) A/C, FIB and I1, that are known for their ability to carry

antimicrobial resistance genes and potential to spread from one bacterium to another. These efforts to study plasmids have involved DNA sequencing and functional assessments of plasmid function, including examining antimicrobial resistance genetics, virulence characteristics and transfer potential. These studies have highlighted that challenges remain in maximizing sequencing data for plasmid analyses, thus ongoing efforts in our research program are focused on the development of a plasmid gene database and improved predictive tools for plasmid characterization based on genome sequencing data. This presentation examines results of our laboratory and bioinformatics approaches to study plasmid genetics, antimicrobial resistance and potential for resistance spread. These findings have contributed to further studies that have shown that microorganisms associated with the gastrointestinal tract contribute to antimicrobial resistance and have the the potential to spread antimicrobial resistance among enteric bacterial populations.

Thomas Hartung, M.D. Ph.D.
Professor and Chair for Evidence-based Toxicology
Johns Hopkins University
Bloomberg School of Public Health
Baltimore, MD, USA

Thomas Hartung, MD PhD, is the Doerenkamp-Zbinden-Chair for Evidence-based Toxicology with a joint appointment for Molecular Microbiology and Immunology at Johns Hopkins Bloomberg School of Public Health, Baltimore. He holds a joint appointment as Professor for Pharmacology and Toxicology at University of Konstanz, Germany; he also is Director of Centers for Alternatives to Animal Testing (CAAT, http://caat.jhsph.edu) of both universities with the portal AltWeb (http://altweb.jhsph.edu). CAAT hosts the secretariat of the Evidence-based Toxicology Collaboration (http://www.ebtox.org), the Good Read-Across Practice Collaboration, the Good Cell Culture Practice Collaboration, the Green Toxicology Collaboration and the Industry Refinement Working Group. As PI, he headed the Human Toxome project (http://humantoxome.com) funded as an NIH Transformative Research Grant. He is the former Head of the European Commission's Center for the Validation of Alternative Methods (ECVAM), Ispra, Italy, and has authored more than 540 scientific publications**.**

Big data and machine learning to predict chemical toxicity

We created earlier a chemical hazard database via natural language processing of dossiers submitted to the European Chemical Agency with approximately 10 000 chemicals and 800,000 associated toxicological studies. We identified repeat OECD guideline tests to establish reproducibility of acute oral and dermal toxicity, eye and skin irritation, mutagenicity and skin sensitization. Based on 350–700 chemicals each, the probability that an OECD guideline animal test would output the same result in a repeat test was 78%–96% (average 81% balanced accuracy, sensitivity 50%–87%, averagage 69%). An expanded database with more than 866 000 chemical properties/hazards was used as training data and to model health hazards and chemical properties. The constructed models automate and extend the read-across method of chemical classification. The novel models called RASARs (read-across structure activity relationship) use binary fingerprints and Jaccard distance to define chemical similarity. A large chemical similarity adjacency matrix is constructed from this similarity metric and is used to derive feature vectors for supervised learning. We show results on 9 health hazards from 2 kinds of RASARs—"Simple" and "Data Fusion". The "Simple" RASAR seeks to duplicate the traditional read-across method, predicting hazard from chemical analogs with known hazard data. The "Data Fusion" RASAR extends this concept by creating large feature vectors from all available property data rather than only the modeled hazard. Simple RASAR models tested in cross-validation achieve 70%–80% balanced accuracies with constraints on tested compounds, i.e. it delivers predictions for about two thirds of chemicals. Cross validation of data fusion RASARs show balanced accuracies in the 80%–95% range across 9 health hazards with no constraints on tested compounds. The approach is continuously being updated and improved and has been implemented as Underwriters Laboratories (UL) Cheminformatics Tool Kit.

Weida Tong, Ph.D.
Director, Division of Bioinformatics and Biostatistics
National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Arkansas, USA

Dr. Weida Tong is Director of Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR/FDA). He has served science advisory board for several multi-institutional projects in Europe and USA. He also holds adjunct appointment at several universities. In addition, he is the founder and board chairperson of newly established international MAQC Society. His division at FDA is to develop bioinformatic methodologies and standards to support FDA research and regulation and to advance regulatory science and personalized medicine. The most visible projects from his group are (1) conducting the Microarray and Sequencing Quality Control (MAQC/SEQC) consortium to develop standard analysis protocols and quality control metrics for emerging technologies to support regulatory science and precision medicine; (2) development of liver toxicity knowledge base (LTKB) for drug safety; (3) *in silico* drug repositioning for the enhanced treatment of rare diseases; and (4) development of various tools such as ArrayTrack™ suite to support FDA review and research on pharmacogenomics. In addition, his group also specializes in molecular modeling and QSARs with specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book chapters.

Of Text and Gene – Analysis of Big Genomics Data with Text Mining Methods

Big data in genomics are diverse and complex. For example, in toxicogenomics, study design often profiles gene expression from assays involving multiple doses and time points, requiring analysis taking this characteristic into specific consideration in toxicity assessment. The genome is often referred to as a book of life: the genome has 30 billion letters (bases), ~25,000 words (genes) comprised by these letters, many sentences/paragraphs (biological processes) that can be constructed with these words to associate with diseases, and these sentences/paragraphs are repeated and spread across 23 chapters (chromosomes). Thus, one can conceptualize a relationship between genes and text; genes and text share many commonalities and characteristics. For example, the same word can appear in different sentences while the same gene can involve in different pathways. Such a commonality suggests that text mining tools could be useful alternatives to analyze genomic data. Topic modeling is a text mining approach, but, by analogy, could be effective in genomics data analysis due to the similar data structure between text and gene dysregulation. In this presentation, we will present the results by applying topic model to a very large toxicogenomics dataset that contains microarray gene expression data from >15,000 samples associated with 131 drugs tested in three different assay platforms (i.e., *in vitro* assay, *in vivo* repeated dose study and *in vivo* single dose experiment) with a design including multiple doses and time points. A set of "topics" (each consists of a set of genes) was determined, by which the varying

sensitivity of three assay systems was observed. Specifically, the drug-dependent effect was more pronounced in the two *in vivo* systems than the *in vitro* system, while the time-dependent effect was reflected the strongest in the *in vitro* system followed by the single dose study and then the repeated dose experiment. The dose-dependent effect was similar across three assay systems. Although the results indicated a challenge to extrapolate the *in vitro* results to the *in vivo* situation, we did notice that, for some drugs but not for all the drugs, the similarity in gene expression patterns was observed across all three assay systems, indicating a possibility of using the *in vitro* systems with a careful design (such as the choice of dose and time point), to replace the *in vivo* testing strategy. The study demonstrated that text mining methodologies such as topic modeling provide an alternative way to other traditional computation means for data reduction in toxicogenomics, enhancing a researcher's capability to interpret the biological information in a reduced data features.

List of Poster Presenters

Poster No.	Presenter's Name	Institution	Abstract Page No.	Category	Title of Presentation
AR-BIC 1	Kaleb Abram	UAMS	68	Graduate	Scientific Computing Facilities at UAMS
AR-BIC 2	Daniel Acheampong	UALR/UAMS	69	Graduate	Bacterial Microbiota profiling in a mock community using Nanopore sequencing: PCR amplification of the full length 16S rRNA gene
AR-BIC 3	Duah Alkam	UAMS	70	Graduate	Deducing Staphylococcus aureus Virulence-Critical Genes
AR-BIC 4	Tanmay Bera	NCTR	71	Post Doc	Image Acquisition and Analysis Techniques for Specie Level Identification of Food Contaminating Beetles
AR-BIC 5	Gehendra Bhattarai	University of Arkansas	72	Graduate	Genome-wide association mapping of downy mildew resistance in spinach germplasm
AR-BIC 6	Zac Campbell	A State	73	Professional	High-throughput Plant Phenotyping at the A-State Phenomics Facility
AR-BIC 7	Tao Chen	NCTR	74	Professional	Evaluation of germline mutations induced by ethyl methanesulfonate in different reproductive development stages of Caenorhabditis elegans using whole genome sequencing
AR-BIC 8	Skylar Connor	UALR/UAMS	75	Graduate	Ribosomal Proteins for Better

					Resolution in Microbial Taxonomy
AR-BIC 9	Celeste Dunn	University of Arkansas	76	Graduate	RNA Sequencing following multiple traumatic brain injuries in a closed-head mouse model
AR-BIC 10	Yheni Dwiningsih	University of Arkansas	77	Graduate	Molecular Genetic Analysis of Drought Resistance and Productivity in US Rice Cultivars
AR-BIC 11	Hong Fang	NCTR	78	Professional	Profiling of Serious Adverse Drug Reactions using FDA-approved Drug Labeling and MedDRA
AR-BIC 12	Bryan Ponce Fuentes	University of Arkansas	79	Graduate	Computational Statistics for Predictive Modeling of Soil pH
AR-BIC 13	Chirag Gupta	University of Arkansas	80	Graduate	The drought tolerance gene classifier of rice
AR-BIC 14	Amer Hasan	University of Arkansas	81	Professional	Role of Gut Integrity and Cellular Tight Junctions in Lameness of Broiler Chickens
AR-BIC 15	PJ Icalia	University of Arkansas	82	Graduate	Genome variability in Prevotella intermedia 17 and other species identified in periodontitis patients and CRISPR array characteristics
AR-BIC 16	Taylor Ingle	NCTR	83	Professional	Profiling of Opioid Information Using FDALabel Database
AR-BIC 17	Donald Johann	UAMS	84	Professional	Scientific and Methodological Advances in Liquid Biopsies to Further the Development of Lung Cancer-Based Precision Medicine
AR-BIC	Sangam Kandel	UALR/UAMS	85	Graduate	Composition of The Gut Microbiome in Delirious and

18					Non-delirious Critically Ill Patients.
AR-BIC 19	Kapil Khadka	NCTR	86	Post Doc	AOP Network as a framework to study adverse effects: a case with DILI
AR-BIC 20	Ting Li	UA Little Rock	87	Graduate	Tracing Drug Induced Liver Injury (DILI) signatures by harnessing cancer cell lines
AR-BIC 21	Zhichao Liu	NCTR	88	Professional	Repurposing Immortalized Cell Line-based Transcriptomic Profiling Assays for Drug-induced Liver Injury with a PRank Method
AR-BIC 22	Karina Medina-Jimenez	A State	89	Post Doc	High throughput kernel phenotyping of corn hybrids grown in Arkansas
AR-BIC 23	David Mery	UAMS	90	Graduate	The Sesquiterpene Lactone Parthenin, from Parthenium hysteropherus, Ablates Acute Myelogenous Leukemia (AML)
AR-BIC 24	Nirman Nepal	A State	91	Post Doc	Mechanisms Underlying the Enhanced Biomass and Abiotic Stress Tolerance Phenotype of Arabidopsis MIOX Over-expressor Line
AR-BIC 25	Ujwani Nukala	UAMS	92	Graduate	Antioxidant Tocols as Potent Radiation Countermeasures
AR-BIC 26	Rajendram Rajnarayanan	A State	93	Professional	Human Computer Interface Guided Approach to Accelerate Drug Discovery
AR-BIC 27	Waltram Ravelombola	University of Arkansas	94	Graduate	Application of Bayesian Models to Investigate the Genetics of Leaf Drought Tolerance in Cowpea at Early Vegetative

					Stage
AR-BIC 28	Suguna Sakkiah	NCTR	95	Professional	Structural Changes Due to Antagonist Binding in Ligand Binding Pocket of Androgen Elucidated by Molecular Docking and Dynamic Simulations
AR-BIC 29	Kate Sanders	Hendrix College	96	Undergraduate	User-friendly programs to identify gene variant information for Medullary Cystic Kidney Disease
AR-BIC 30	Abdulkarim Shwani	University of Arkansas	97	Professional	Directed Genome Evolution for Macrophage-killing in Staphylococcus agnetis an Agent of Bacterial Chondronecrosis with Osteomyelitis in Broilers.
AR-BIC 31	Visanu Wanchai	UAMS	98	Graduate	Using 16S rRNA for Microbial Classification: A Cautionary Tale
AR-BIC 32	Xiaofan Wang	University of Arkansas	99	Graduate	Lifelong dynamics of the swine gut microbiome: from birth to market
AR-BIC 33	Charity Washam	UAMS	100	Professional	Discontinuation of trichloroethylene attenuated the DNA methylation changes in polycomb protein binding sites in effector/memory CD4+ T cells observed after continuous developmental exposure
AR-BIC 34	Matthew Windsor	UAMS	101	Graduate	
AR-BIC 35	Yifan Zhang	UA Little Rock	102	Graduate	A pan-cancer study analysis reveals novel regulations of

					LRP1B in cancer pathways
AR-BIC 36	Bazgha Umair Zia	University of Arkansas	103	Professional	Evaluation and Association Analysis of Downy Mildew Resistance in USDA Spinach Germplasm

AR-BIC 1

Scientific Computing Facilities at UAMS

Kaleb Abrahm, B.S.[1], Lawrence Tarbox, Ph.D.[1]
[1] University of Arkansas for Medical Sciences, Department of Biomedical Informatics, Center for High Performance Computing

The University of Arkansas for Medical Sciences (UAMS) College of Medicine invested in a High-Performance Computing Cluster and associated storage services. Low priority batch computing time on the 200+ node Beowulf cluster is available to researchers without charge. Over 6 PB of data storage is available to researchers, divided between 2 PB of high speed (20 GB/s) cluster storage used primarily for scratch memory, and 4+ PB of object storage (shared cost model) for data archiving. The cluster includes 96 nodes with Xeon Phi many-core processors (64 cores per node), the majority with 384 GiB of memory (16 nodes with 192 GiB of memory), giving a total of 6,144 Phi cores. Another 96 nodes are dual socket regular Xeon processor, 28 cores per node (2,288 total cores), with 128 GiB memory per node. There are 3 dual Xeon socket nodes with 24 cores and 128 GiB memory each that also include dual Nvidia P100 Tesla GPU cards. The compute nodes are interconnected with a 100 GB/s Omnipath high speed, low latency network supporting inter-process communications via MPI and other libraries. This network is also used by the fast cluster storage. Cluster activities and logins are managed and monitored by 5 support nodes using the Viewpoint/Moab/Torque cluster software.

We are experimenting with bridges between the UAMS High Performance Computing and the High-Performance Computing facilities at the University of Arkansas in Fayetteville (UAF). The two centers compliment each other in terms of their capabilities, and researchers may utilize the resources at either center.

We are also experimenting with web-based access to the computing facilities for researchers who do not have direct campus access. We are creating a “Science DMZ”, to reduce the barriers for accessing the systems, which we hope will extend across the state and beyond.

AR-BIC 2

Bacterial Microbiota profiling in a mock community using Nanopore sequencing: PCR amplification of the full length 16S rRNA gene

Daniel A. Acheampong[2], Piroon Jenjaroenpun[1], Wongsurawat Thidathip[1], Intawat Nookaew[1]

[1]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR
[2]Department of Bioinformatics, University of Arkansas at Little Rock, AR

Technological advancements in genome sequencing have led to a tremendous improvement in genomic research whiles reducing the time and cost in sequencing. With the emergence of the third-generation single molecule sequencing technologies such as MinION (Oxford Nanopore Technologies (ONT)), the full length of the 16S ribosomal RNA (rRNA) (~1,500bp) can be sequenced rapidly within a short period of time as compared to other sequencers. In this study, we aim to assess the performance of the MinION in assigning taxonomy at the species level using the full length 16S rRNA. We sequenced two bacterial mock communities (ATCC MSA-1003™ & ATCC MSA-1002™) on MinION sequencer using a direct PCR 16S barcoding library kit and the sequenced data were analyzed using a published Bioinformatics workflow. We retrieve all the microbiota composition present in the mock communities with the exception of the *Bifidobacterium adolescentis*. The poor detection of the *Bifidobacterium adolescentis* in the mock was due to the choice of primers used (27F and 1492R) which failed in the amplification of the species. The MinION performs well in retrieving most of the microbiota despite the high error rate of ONT reads. Taxonomic identification with MinION at the clinics will aid in the rapid characterization of pathogens and development of antibiotics.

AR-BIC 3

Deducing *Staphylococcus aureus* Virulence-Critical Genes

Duah Alkam[1], Zulema Udaondo[1], Kaleb Abram[1], David Ussery[1]
[1-]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR

Staphylococcus aureus is a prominent bacterial pathogen that causes a wide range of severe infections such as endocarditis and osteomyelitis. Growing antibiotic resistance among other factors limits our ability to treat *S. aureus* infections. Methicillin-resistant S. aureus (MRSA) strains lead to increasingly common infections in the hospital and community acquired settings. These are difficult to treat due to their resistance to many antibiotics. Therefore, it is of paramount importance to understand the genetic factors and molecular mechanisms that enable *S. aureus* to escape antibiotics. We took a comparative-genomics approach to elucidate genes that are crucial to MRSA strains. We identified 12 phylotypes based on genome K-mer analysis of more than 10,000 *S. aureus* genomes from GenBank. We then identified and compared the core genomes of each of the phylotypes. We expect to characterize genes unique to the MRSA strains. These may play a crucial role in exacerbating the virulence of these strains.

AR-BIC 4

Image Acquisition and Analysis Techniques for Specie Level Identification of Food Contaminating Beetles

Tanmay Bera[1], Hongjian Ding[2], Howard Semey[2], Leihong Wu[1], Zhichao Liu[1], Amy Barnes[2], Darryl Langley[2], Monica Pava-Ripoll[3], Himansu Vyas[2], Weida Tong[1], Joshua Xu[1*]
[1]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079.
[2]Food Chemistry Laboratory , Arkansas Laboratory, Office of Regulatory Affairs, US Food and Drug Administration, Jefferson, AR.
[3]Office for Food Safety, Center for Food Safety and Applied Nutrition, US Food and Drug Administration, College Park, MD 20740.

Insect pests such as pantry beetles are one of the major reasons for the contamination and wastage of food products. The severity of such damages varies from one species to another, so their species level identification is necessary for better managing contamination scenarios. In this context, we have been developing automated machine learning methods that are more efficient than the current method of manual micro-analysis. Our initial studies have shown that food contaminating beetles could be identified by recognizing the patterns on their elytron (hardened forewings) which is unique for every species. Acquiring better quality image that reveal the finest details of the elytra patterns was found crucial in achieving accurate specie level identification. In this study we explored various optical settings (such as optical filters, lighting conditions, magnifications etc.) to obtain a standard imaging condition that can consistently produce good quality high resolution (HR) images of beetle elytra across various species of pantry beetles. We observed the optimized imaging set-up did result in obtaining distinct images of elytra for 6 most difficult (from our previous studies) to identify species of pantry beetles. They were also detailed enough to allow distinguishing one species from another, even for beetles from the same genus. Using this set-up, we have created an HR image database for 15 different species of pantry beetles which would be extended to 15 more species. Work is currently underway to develop an efficient pattern recognition algorithm that can yield high species level accuracy for an extensive array of beetle species. In summary, our work demonstrate that a combination of image acquisition and image processing technique is the key in achieving better accuracy in identifying the species of food contaminating beetles, which may eventually help better manage scenarios of food contaminations due insect pests in the future.

Disclaimer:
The views expressed in this work are those of the authors only and do not necessarily express the views/policies of the U.S. FDA.

AR-BIC 5

Genome-wide association mapping of downy mildew resistance in spinach germplasm

Gehendra Bhattarai[1], Ainong Shi[1], Chunda Feng[2], Braham Dhillon[2], James C. Correll[2], Beiquan Mou[3]
[1]Dept. of Horticulture, University of Arkansas, Fayetteville, AR 72701
[2]Dept. Plant Pathology, University of Arkansas, Fayetteville, AR 72701
[3]Crop Improvement and Protection Research Unit, USDA-ARS, Salinas, CA 93905

Spinach (*Spinacia oleracea*), an important leafy vegetable crop, showed a significant increase in consumption in the U.S in the last two decades. Downy mildew (DM) caused by an obligate oomycete *Peronospora effusa* [=*P. farinosa f. sp. spinaciae* (Pfs)] is the most important disease affecting fresh-market spinach production in California, Arizona, and in Europe. Seventeen different races of Pfs have been reported in spinach and many of these races (> 10) were identified in the last two decades. Rapid emergence of new Pfs race and regular breakdown of the resistance genes deployed in the major cultivars indicates an urgent need to screen wide germplasms and to identify new resistance sources. Utilization of host genetic resistance is the most practical and economical disease management practice. More than 400 spinach genotypes collected from a wide geographical region and maintained at NCRPIS were evaluated at the USDA research station, Salinas, CA and Yuma agricultural center, the University of Arizona in 2018 and 2019. The purpose of the field evaluation was to identify field resistance to the downy mildew pathogen under natural disease pressure under field conditions and to identify resistance-governing QTLs. The spinach accessions evaluated were originated from 37 countries and disease severity was rated on a scale of 0-100% based on a percentage of leaf area showing signs or symptoms of downy mildew. A wide variation in downy mildew disease severity was observed among the evaluated spinach genotypes and GBS generated SNPs were used to conduct an association analysis. Whole-genome resequencing of all evaluated spinach genotypes is underway, and we plan to use the SNPs marker identified from the population resequencing to conduct the genome-wide association analysis. Identification of markers associated with the resistance allele allow multiple resistance gene pyramiding, and to deliver more durable downy mildew resistant spinach cultivars.

AR-BIC 6

High-throughput Plant Phenotyping at the A-State Phenomics Facility

Zachary Campbell[1], Lucia Acosta-Gamboa[1], Nirman Nepal[1], Shannon Cunningham[1], and Argelia Lorence[1,2]
[1] Arkansas Biosciences Institute, Arkansas State University; [2] Department of Chemistry and Physics, Arkansas State University, P.O. Box 419, State University, AR 72467, USA. Email: alorence@astate.edu

The A-State Phenomics Facility offers a variety of high-throughput plant phenotyping assay capabilities, including multi-well plates for seed phenotyping; petri dishes to assess *in vitro* cultures, seed germination, and seedling growth and vigor; tray configurations for small plants such as Arabidopsis and tobacco; and pot configurations for larger, faster growing plants, such as rice, maize, soybean, and tomato. We routinely test plants growing in both soil and hydroponics, and a number of stress tolerance protocols have been optimized, including those to assess water limitation, heat, nutrients, light, and salinity. We continue to update our technology and protocols in order to achieve the highest level of excellence during each experiment. Utilizing visible, fluorescence, near and far infrared sensors, and proprietary and open source algorithms, we can obtain a wealth of readouts to quantify plant size, color, architecture, leaf temperature, and overall health to empower plant biology research.

AR-BIC 7

Evaluation of germline mutations induced by ethyl methanesulfonate in different reproductive development stages of Caenorhabditis elegans using whole genome sequencing

Tao Chen[1], Bohu Pan[1], Jian Yan[1], Ying Chen[1] Timothy W. Robison[2], Mark Powley[2], Robert H. Heflich[1] and Hua Du[3]
[1]Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR; [2]Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD and 3Department of Chemistry, University of California at Riverside, Riverside, CA

Germline mutation is a heritable change in the DNA that occurred in a germ cell and can be passed on to their descendants. Germline mutations are responsible for more than 5000 known inherited diseases. A practical assay for detecting germline mutations is desired for regulatory purpose. Caenorhabditis elegans (C.elegans) has a short hermaphroditic life cycle, a relatively small genome, and a large number of offspring, making it an ideal model for assaying germline mutagenicity of chemicals. In this study, we evaluated the responses of different germline development stages of the worm to mutagenic insults to assess whether the model can appropriately detect chemical-induced mutations in male and female germs. Parent worms were treated with 25 mM ethyl methanesulfonate (EMS), a model germline mutagen, for 4 hours in L3 larval stage when the germ line undergo robust proliferation, L4 larval stage when the hermaphrodite develops spermatogenesis, and young adult stage when the hermaphrodite ceases spermatogenesis and switches to oogenesis. The whole genome of F1 offspring from each individual worm treated with EMS or the vehicle was sequenced and the sequencing data were analyzed with a bioinformatics pipeline including VarScan, a variant calling program. The mutation frequencies (MFs) were 185-fold, 130-fold and 175-fold higher in the L3, L4 and young adult treatment groups, respectively, than the control group (MF is 2 ×10-8). More than 90% of the induced mutations are G:C >A:T transitions, a signature mutation type induced by EMS. The MFs induced in sperm germs were slightly lower than those in early germ cell proliferation stage and in egg germs. The results suggest that C. elegans can sensitively detect exposure of germline mutagens in both male and female germ cells, as well as in early germline proliferation stage.

AR-BIC 8

Ribosomal Proteins for Better Resolution in Microbial Taxonomy

Skylar Connor[1,2], Visanu Wanchai[2], Michael Robeson[2], Trudy Wassenaar[2,3], David Ussery[2]

[1]Joint UALR/UAMS Bioinformatics Program, University of Arkansas at Little Rock, Little Rock, Arkansas, 72204
[2]Department of Biomedical Sciences, University of Arkansas for Medical Sciences, Little Rock, Arkansas, 72205
[3]Molecular Microbiology and Genomics Consultants, Zotzenheim, Germany

Background: 16S ribosomal RNA (rRNA) is the known standard for taxonomic classification within the microbial community. Often the sequence reads are too short to obtain adequate full length16S rRNA sequences. However, even when the full 16S rRNA sequences are present and used to build microbial trees, they only show species level information. The resolution is unable to differentiate multiple genomes within the same species; for example, a pathogenic Escherichia coli O157:H7 and Shigella or a commensal strains of E. coli and a probiotic strain all cluster as the same species. Furthermore, with the level of information provided by the standard method of 16S rRNA sequencing it is difficult for one to correctly classify rapidly growing organisms, like Vibrio, which have multiple divergent copies of their 16S rRNA. Classification of these genomes end up being dependent on which one of the many different 16S sequences are chosen.
Results: More than 10,000 bacterial genomes in the NCBI database do not have full length 16S rRNA sequences; this is due to the resolution obtained from second generation sequence reads. On average each bacterial genome contains a full complement of about 50 ribosomal protein sequences. The genomes from bacterial type strains were used for comparison of genomes from closely related species; we find agreement with known taxonomy, and further, we can get strain level separation using ribosomal proteins.
Conclusion: We find that ribosomal proteins can be used to build phylogenetic trees for bacterial genomes; using high throughput computing this can be extended to more than a hundred thousand genomes currently available. Ribosomal protein trees can give strain-level resolution for bacterial genomes within the same species.

AR-BIC 9

RNA Sequencing following multiple traumatic brain injuries in a closed-head mouse model

Celeste Dunn[1], Sara Venier[2], Jeff Wolchok[3], and Kartik Balachandran[3]
[1]Cell and Molecular Biology Program, University of Arkansas, Fayetteville, AR 72701, USA
[2]Department of Biology, University of Arkansas, Fayetteville, AR 72701, USA
[3]Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR 72701, USA

Traumatic brain injuries (TBI) are the cause of a silent epidemic in today's society. While the pathophysiology following a single TBI is not fully understood, even less is known about the response following multiple TBIs. Through a closed-head mouse model, we studied the effects of both single and double TBIs at both a mild (<90 g forces) and moderate (>120g force) severity. We employed the use of RNA Sequencing to study the alteration in the genome of mouse cortical tissue after TBI. Our findings demonstrate a correlation between number of impacts and up or down regulation of genes known to be associated with the blood-brain barrier, inflammation, the extracellular matrix, astrogliosis, and the wnt and beta-catenin signaling pathways.

AR-BIC 10

Molecular Genetic Analysis of Drought Resistance and Productivity in US Rice Cultivars

Yheni Dwiningsih[1], Anuj Kumar[1], Julie Thomas[1], Sara Yingling[1] and Andy Pereira[1]
[1]Department of Crop, Soil and Environmental Sciences, University of Arkansas, Fayetteville, AR, USA
Corresponding author: apereira@uark.edu

Rice is the staple food for half of the world population, with USA as the third largest exporter of rice, with an export value of US$ 1.8 billion. Rice also uses 2-3 times the water as other food crops, which totals 30% of the world's freshwater resources world-wide. Stability of rice production is facilitated by economic use of water, which is most essential during flowering and grain formation. In our research we screened adapted US rice cultivars, comprising tropical japonica rice genotypes, for drought resistant (DR) and water use efficiency (WUE) traits to search sources for breeding US rice cultivars for a water saving agricultural system. A RIL population derived from varieties Kaybonnet (DR) and ZHE733 (sensitive), termed K/Z RILs, available from the USDA Dale Bumpers National Rice Research Center and chosen for genetic analysis of DR/WUE traits. The RIL population was screened in Fayetteville (AR) by controlled drought stress treatment at flowering, and the effect of stress was quantified by counting number of filled grains per panicle. Based on the DR scores, a primary genetic screen was done using bulked segregant analysis (BSA), where sets of 10 DR and sensitive RIL plants were used for screening of markers to find polymorphisms linked to the yield-related trait number of filled grains per panicle under drought. From this BSA screen, a total of 6 polymorphic markers were identified, with most of the SSRs displaying a single allele fragment of 100-400 bp size. The SSR markers with potential linkage to drought resistance are RM9, RM109, RM236, RM114, RM131 and RM139. Presently, additional methods are being used to map the DR and generate markers for screening and selection, to provide a foundation for molecular breeding of drought resistance in US rice.

AR-BIC 11

Profiling of Serious Adverse Drug Reactions using FDA-approved Drug Labeling and MedDRA

Hong Fang[1], Leihong Wu[1], Taylor Ingle[1], Zhichao Liu[1], Shraddha Thakkar[1], Anna Zhao-Wong[2], Steve Harris[1], Guangxu Zhou[1], Junshuang Yang[1], Joshua Xu[1], Weida Tong[1]
[1]National Center for Toxicological Research, U.S. FDA, Jefferson, Arkansas 72079; [2]MedDRA MSSO, 7575 Colshire Dr., McLean VA USA 22102

Background: Adverse Drug Reactions (ADRs) contribute to over 100,000 deaths per year in the US. FDA-approved drug labeling contains rich ADR information collected from clinical trials and post-market surveillance, to promote the safe use of pharmaceuticals and can be an important source for the study of ADRs. Currently, ADR studies underutilize the drug labeling information and medical standard terminology. There are three ADR related sections in drug labeling, namely BOXED WARNING (BW), WARNINGS AND PRECAUTIONS (WP), and ADVERSE REACTIONS (AR). In this study, Medical Dictionary for Regulatory Activities (MedDRA) standard terms were used for data mining of the ADR sections via Oracle text search, with specific focus on profiling serious adverse drug reactions (sADRs) from BW.
Result: 1164 FDA-approved drug labeling documents were selected from the FDALabel database and included all current single-ingredient human prescription drugs approved by the FDA as a New Drug Application. The ADRs in BW, WP, and AR were identified by MedDRA terminology and extracted by Oracle text query. We compared the top 20 MedDRA Preferred Terms (PTs) among BW, WP, and AR, and found that six PTs (Death, Pregnancy, Depression, Hemorrhage, Cardiac Failure, Infection) overlapped between BW and WP. We also found that sADRs were prevalent in System Organ Classes (SOCs) such as nervous system disorders, psychiatric disorders, cardiac disorders, and hepatobiliary disorders. Furthermore, Hierarchical Cluster Analysis (HCA) revealed that drugs within the same therapeutic categories might be associated with similar sADRs (e.g., nervous system drug class was found to be highly associated with drug abuse terms such as dependence, substance abuse, and completed suicide).
Conclusion: MedDRA standard terminology combined data mining techniques can enhance the analytical abilities in uncovering information from drug labeling. The proposed bioinformatics approach is valuable in supporting robust and consistent adverse event monitoring, drug safety research, and for the advancement of pharmacovigilance.

AR-BIC 12

Computational Statistics for Predictive Modeling of Soil pH

Bryan A. Fuentes[1], Minerva J. Dorantes[1]; Phillip R. Owens[2]; David Miller[1]
[1]CSES-University of Arkansas, Fayetteville-AR, United States, [2]USDA-ARS, Dale Bumpers Small Farms Research Center, Booneville-AR, United States.

Comprehension of physical, chemical and biological properties of soils is fundamental to establish adequate management practices involving not only the soil system, but also water, microorganisms and crops. The prediction of soil properties through statistical models becomes useful when precise management is required, but the lack of resources does not allow intensive soil sampling campaigns. Advances in distributed computing have allowed the application of computationally intensive, but robust statistical analysis. Data mining, feature engineering and predictive modeling are some of the computational statistics procedures now commonly applied to large, non-homogeneous environmental datasets. This study aimed to evaluate the performance of 3 different schemes for feature engineering and predictive modeling of soil pH. Soil pH data consisted of 663 soil surface samples collected in the study site, in El Salvador. Three schemes were established to select the most accurate feature engineering procedure and predictive model for soil pH. The first scheme consisted of feature selection based on rotated principal component analysis (PCA), coupled with a rule-based inference. The second scheme consisted on recursive feature elimination (RFE), coupled with a random forest regression. The third scheme applied an elastic net regularization to select significant features within a generalized linear model (GLM). In all schemes, an initial set of 23 predictors were considered. The initial 23 predictors meant to capture the influence of topography, climate and vegetation on soil properties, specifically pH. The statistical performance of the 3 schemes was evaluated in terms of mean absolute error (MAE), root mean squared error (RMSE) and agreement coefficient (AC). Results show that the RFE/random forest regression scheme outperformed both the rotated PCA/rule-based inference and the elastic net/GLM schemes. Overall, the adequate selection of predictive features via statistical methods allowed parsimonious models which provided robust predictions of soil pH.

AR-BIC 13

The drought tolerance gene classifier of rice

Chirag Gupta[1], Julie Thomas[1], Anuj Kumar[1] and Andy Pereira[1]
[1]Department of Crop, Soil and Environmental Sciences, University of Arkansas, Fayetteville, AR, USA.

With the generation and availability of large functional genomics datasets, associated with differences in a trait, there is a need to identify genes that can be found associated with the trait phenotype(s). We present here, a systematic method to prioritize trait-relevant genes from large-scale gene expression data. A binary classifier was trained to discriminate between network connectivity patterns of known 'drought tolerance' (DT) genes (positive labels) in rice, from those that evidently do to not associate with drought (negative labels). Our method generated a rank for each gene according to its predicted association to DT. We estimated the accuracy of our method using five-fold cross validation, as well as independent rank enrichment tests with publicly available drought GWAS predicted loci not used in training. These evaluations show that our method generally outperforms other approaches in gene prioritization, and is robust to noise in the training data. To facilitate implementation of our method to other organisms and traits, we developed a docker application packaged with all our codes and software dependencies required to execute all the aspects of our proposed method, including base network building and the machine learning based prediction module. Besides this standalone executable application, an interactive web interface is also conceptualized. We anticipate that our proposed method and accompanying software will greatly benefit researchers, who wish to select a few actionable candidate genes, using expression datasets relevant to any trait for which substantial amount of prior information is available.

AR-BIC 14

Role of Gut Integrity and Cellular Tight Junctions in Lameness of Broiler Chickens

Amer Hasan, Douglas D. Rhoads, Adnan Alrubaye
University of Arkansas, Cell and Molecular Biology Program

The goal of this study is to identify the role of tight junction and adherens proteins as relates to the integrity of gut linings of broiler chickens on wire and litter floors. Lameness is one of the main metabolic diseases linked to fast growth in broilers. It is a significant problem in the poultry industry, resulting in hundreds of millions of dollars in lost revenue annually. In commercial broilers, the most common cause of lameness is bacterial chondronecrosis with osteomyelitis (BCO). Our hypothesis is that bacteria cross the epithelial lining of the gut and respiratory tract into the bloodstream and eventually colonize the growth plate of rapidly growing long bones in commercial broiler chickens. Our lab has embarked on numerous projects to characterize the genomic features of bacteria implicated in the disease; our lab has also worked to understand the physiology and histopathology of guts of chickens grown on litter or wire flooring. Previous work has demonstrated that growth on wire flooring, dramatically increases the incidence of BCO. Our hypothesis is that the wire flooring compromises gut integrity, allowing bacteria to translocate into the blood stream. Our immediate is goal is to curb lameness, and as such, we have administered different supplements, probiotics and organic acids to broiler birds on both floor types, with some positive results. We are quantifying the levels of expression of tight junction and adherens proteins (between the treatments) along with immunohistopathology to obtain a better understanding of the role of gut integrity in BCO lameness.

AR-BIC 15

Genome variability in Prevotella intermedia 17 and other species identified in periodontitis patients and CRISPR array characteristics

Peter James C. Icalia[1], Vibha Srivastava[1], and Sherwin Reyes[2]
[1,2]Cell and Molecular Biology Program, University of Arkansas, Fayetteville, Arkansas, USA
[2]Department of Crop, Soil and Environmental Science, University of Arkansas, Fayetteville, Arkansas, USA
[3]Department Biochemistry and Molecular Biology, Miller School of Medicine, University of Miami, Miami, Florida,

The relationship of Prevotella intermedia 17 with other species and strains associated with periodontitis was analyzed, orthologs and unique genes were identified and Open Reading Frames and CRISPR arrays were characterized using the downloaded sequences from NCBI.
Prevotella intermedia 17 showed a closer genetic relationship to Fusobacterium periodonticum 2131 (Fs_2_1_31), Porphyromonas gingivalis ATCC33277 (Pg_ATCC33277), and Tannerella forsythia ATCC43307 (Tf_ATCC43037) after clustering together under the same clade but branched out to an arm distinct from the rest of the three species. The GC content of Prevotella intermedia 17 is close to the other Prevotella species and the CDS or genome size were positively correlated and can be predicted using the regression equation y = 585.58x + 658.9.
Orthologs shared by selected Prevotella species were associated to metalloendopeptidase activities, integrators of pattern recognition and signaling factors for immunity. There were CRISPR arrays in both chromosome of Prevotella intermedia 17 but only spacers in chromosome 2 showed similarity to some bacteriophage and other bacterial genomes.

AR-BIC 16

Profiling of Opioid Information Using FDALabel Database

Taylor Ingle[1], Zhichao Liu[1], Darshan Mehta[1], Shraddha Thakkar[1], Stephen C Harris[1], Junshuang Yang[1], Guangxu Zhou[1], Leihong Wu[1], Catherine Li[2], Crystal Allard[2], Joe Meehan[1], Joshua Xu[1], Lilliam Rosario[2], Weida Tong[1], and Hong Fang[1]
[1]National Center for Toxicological Research, U.S. Food & Drug Administration (FDA), Jefferson, AR 72079, USA; 2Office of Computational Science (OCS), Center for Drug Evaluation and Research (CDER), U.S. FDA, Silver Spring, Maryland 20993, USA

The opioid epidemic has drawn national attention and has been officially declared a state of emergency in the US. ~100 million Americans experience chronic pain resulting in an estimated $560-635 billion/year economic burden . As public awareness and education improve, the scope of this epidemic continues to grow indicating that more information and action is needed. FDA-approved drug labeling contains a complete set of opioid drug products and can be used as a reliable source to aid in understanding the mechanisms and characteristics of opioids. FDALabel Database is a publicly available web application that allows for customized searches against the entire text of drug labeling. To demonstrate FDALabel's utility, a case-study was designed to collect information from opioid drug products. Queries were performed based on opioid-related pharmacologic classes, chemical structures, DEA schedules, MedDRA terms, labeling sections, and drug interactions. Using this approach, 12 Pharmacologic Drug Classes and 32 active ingredients were identified. The identified drugs are indicated for analgesic, gastrointestinal, addiction, or opioid overdose treatment. Of the 32 drugs, 24 have Boxed Warnings (BW), 22 are controlled substances, 14 are classified as CII (high potential for abuse), and 11 are in the Full Opioid Agonist pharmacologic class. Twenty-seven different instances of drug-drug interactions were also identified. The top 10 occurring MedDRA preferred terms in BW, which correlate to adverse reactions, are Death, Overdose, Dependence, Substance abuse, Depression, Respiratory depression, Nervousness, Coma, Withdrawal syndrome, and Drug interaction. The ever-growing epidemic of opioid abuse is of great interest to the FDA, which is committed to advancing efforts to address the crisis of misuse and abuse through awareness, education, and research. Here, we provide a comprehensive overview of current opioid information using FDA-approved drug product labeling to support drug safety research, and for the advancement of opioid abuse & addiction prevention.

AR-BIC 17

Scientific and Methodological Advances in Liquid Biopsies to Further the Development of Lung Cancer-Based Precision Medicine

Donald J. Johann, Jr.[1], Mathew Steliga[1], Jason Meusse[1], Erich Peterson[1], Ik Jae Shin[1], Susanne Jeffus[1], Sanjay Maraboyina[1], Nikhil Meena[1], Mary Yang[2], Karl Walker[3], Andy Pereira[4], Huixiao Hong[5], Jason Liem[1], Meei Liu[1], Joshua Xu[5], Weida Tong[5]

[1]University of Arkansas for Medical Sciences, Little Rock, AR, [2]University of Arkansas at Little Rock, Little Rock, AR, 3University of Arkansas at Pine Bluff, Pine Bluff, AR, [4]University of Arkansas, Fayetteville, AR, [5]National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR

BACKGROUND: Blood-based testing for cell free DNA (cfDNA) for oncology applications, and particularly lung cancer, are under rapid scientific development and regulatory evaluation. The clinical implications are enormous with promises to change the practice of oncology. Liquid biopsy methodologies allow for plasma genotyping of solid tumors via cfDNA, and are an application of deep next generation sequencing (NGS). These assays are rapidly entering the clinical domain of research-based monitoring in translational oncology, especially for thoracic malignancies, where they offer a much safer alternative to traditional tissue-based biopsies. Potential applications for these cfDNA assays include: i) initial diagnosis, ii) response to therapy and follow-up, iii) tumor evolution, and iv) minimal residual disease evaluation. These advanced molecular diagnostic assays will greatly contribute towards the goals of precision cancer medicine, and especially regarding treatment decisions in the adjuvant setting, where avoiding over-treatment and unnecessary toxicity are a prime objective.
The use of advanced molecular profiling approaches on individual patient tumor tissues to arrive at rationally-derived therapy decisions for a patient's cancer (vs. categorical therapy assignments) are now being pursued in several advanced clinical trials. Delivering this "right" cancer treatment to the right patient at the right dose and the right time, is a next logical step and a seminal aspect of adjuvant therapy decision making.
RESULTS: The UAMS IRB has approved the clinical protocol and patient enrollment is in progress. An assortment of liquid biopsy methodologies are in development. A variety of novel and complex bioinformatic approaches, designed to address the needs of the clinical trial, are developed and continue to be enhanced. Solid tumor material and cfDNA from routine blood draws are being subjected to molecular profiling with NGS and ddPCR with results being compared, contrasted, and integrated. Model systems development utilizing patient derived xenografts (PDX) are in progress.
CONCLUSIONS: In order to improve outcomes for cancer patients, there is an urgent need for advanced clinical trials utilizing: i) innovative assays, and ii) molecular profiling with cutting-edge bioinformatics, and iii) advanced model systems such as PDX. This is all happening in Arkansas.

AR-BIC 18

Composition of The Gut Microbiome in Delirious and Non-delirious Critically Ill Patients.

Sangam Kandel[1,2], Thidathip Wongsurawat[2], Piroon Jenjaroenpun[2], Intawat Nookaew[2], Michael S. Robeson II[2], David Ussery[2], Anne Sofie Andreasen [2,3]
[1]Department of Bioinformatics, University of Arkansas at Little Rock, Little Rock, AR 72205
[2]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR 72205
[3]Department of Anesthesiology and Intensive Care, University Hospital Herlev, Copenhagen, Denmark

Background: Delirium is defined as a disturbance of consciousness with inattention, accompanied by a change in cognition observed among critically ill patients. About 2 out of 3 patients in ICU get delirium. Previous microbiome studies have shown dynamic changes in gut microbiota of ICU patients in the acute phase of critical illness. But there are no any reported studies about the changes in the composition of the gut microbiota during the development of delirium in critical illness. Since the knowledge on pathogenesis, prevention and treatment of the condition is still limited, an understanding of the association of changes in gut microbiota and critical illness and related conditions could eventually lead to new prevention and treatments strategies. The objective of this study is to investigate if there is an alteration in the gut microbiome among delirious and non-delirious patients admitted in hospital ICU.
Methods: Rectal swab samples were collected from 21 delirious and 19 non-delirious patients admitted at University Hospital in Herlev, Denmark. Ten samples of the nurses working in ICU were used as positive controls. All the patients with dementia, delirium and age below 18 at the time of admission were not included in the study. DNA was extracted from the samples and 16S rDNA sequencing was performed in Illumina Hi-Seq platform. Taxonomic classification and microbial diversity were analyzed using QIIME2.
Results and Conclusion: A total of 42 samples (control-9, delirious-17, and non-delirious-16) were used for diversity analysis after quality control. Preliminary results suggested a difference in the microbial diversity among the healthy control groups and diseased individuals. Data analysis is being performed to compare the microbiome change in delirious and non-delirious patients. Further analysis could be done to find gut microbiota-derived signature among delirious.

AR-BIC 19

AOP Network as a framework to study adverse effects: a case with DILI

Kapil K. Khadka[1], Dong Wang[1]

Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, Jefferson, AR, 72079

Due to the well-known limitations for current animal testing based approaches to predict drug induced liver injury (DILI), there is heightened interest in incorporating high throughput assays into the evaluation framework for DILI risk. However, the diverse and high dimensional nature of these data poses serious challenges for common data mining and machine learning techniques even with recent advances. Integrating high throughput assay information with mechanistic knowledge in the form of expert opinion and literature findings might provide a promising approach to fully utilize the power of both mechanistic understanding and new testing technologies. In this presentation, we discuss our pilot study using adverse outcome pathway (AOP) networks to provide a base model for incorporating high throughput data from L1000, CMap, and Tox21 for gene expression changes and nuclear receptor binding. AOP networks were formed by integrating published AOPs for liver steatosis, cholestasis, fibrosis, and liver tumor. Information for relevant nuclear receptors and genes was then extracted from these networks. We obtained measurements on nuclear receptor binding and differential gene expression for a collection of drugs in the Liver Toxicity Knowledge Base. A rule ensemble learning model was then built to infer liver toxicity from these molecular predictors resulting in competitive performance with other approaches. Our result suggests that current knowledge encoded in AOPs can be successfully utilized for dimension reduction for high throughput data and leading to capable predictive models. With continued improvement in AOP development and new testing technologies, combining mechanistic insight with high throughput data holds great promise in advancing DILI risk assessment.

AR-BIC 20

Tracing Drug Induced Liver Injury (DILI) signatures by harnessing cancer cell lines

Ting Li[1,2], Shraddha Thakkar[1] and Weida Tong[1]
[1]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA
[2]Bioinformatics, Joint program of University of Arkansas at Little Rock and University of Arkansas for Medical Sciences, Little Rock, AR, USA

Background: Drug safety assessment is one of the primary challenges for drug development and for regulatory applications. Because the existing preclinical models have limitations in their ability to predict DILI, there has been significant efforts to develop alternative approaches for predicting DILI. Some of these approaches have emerged as critical tools in regulatory decision-making initiatives, for example the EU's REACH (Replacement, Reduction and Refinement of animal use). 21st century toxicology relies heavily on high throughput technologies such as genomics, and it is important to foster the inclusion of genomic data to improve DILI prediction approaches. This study aims to compare and evaluate the potential of genomics-based data to predict DILI. In this study we used data from the CMAP dataset for two different cancer cell lines (MCF7 and PC3) for 276 drug compounds.
Method and Results: In the data preprocessing, we applied two microarray normalization methods, MAS5 and RMA. We used consensus modeling approach by applying three different machine learning models (Logistic regression, Linear Discriminant Analysis, Gradient boosting) to predict DILI signatures using whole genome array data generated using cMAP. We got the following results. The accuracy, sensitivity and specificity on test dataset were 67.38%, 90.32% and 18.8%; while the accuracy, sensitivity and specificity on validation dataset were 75.58%, 89.55% and 26.32%.
Conclusion: This study was performed to generate evidence to support establishing realistic expectations for the use of genomic technologies for DILI prediction and predictability of DILI signatures using cancer cell lines. We plan to explore further the understanding of DILI signature predictability in difference chemical class and therapeutic category.

AR-BIC 21

Repurposing Immortalized Cell Line-based Transcriptomic Profiling Assays for Drug-induced Liver Injury with a PRank Method

Zhichao Liu[1*], Liyuan Zhu[1], Shraddha Thakkar[1], Ruth Roberts[2,3], Weida Tong[1]
[1] National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas 72079, United States
[2] ApconiX, BioHub at Alderley Park, Alderley Edge, SK10 4TG, UK
[3] University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

In vitro Toxicogenomics (TGx) as an alternative means to animal studies holds great promise in risk assessment. However, TGx studies often suffer from limited sample size and few available cell types. To investigate the potential of repurposing accumulative transcriptomic profiles data of immortalized cell lines in risk assessment, we carried out a comprehensive assessment of the transferability between transcriptomic profiles from three cancer lines (HL60, MCF7, and PC3). These three lines were compared in Connectivity Map (CMap) and in toxicogenomic datasets (human primary hepatocytes and rat in vivo repeated dose) from the Open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATEs) database using our developed Pair Ranking (PRank) method. A moderated concordance was observed in HL60 versus human primary hepatocytes (PRank score = 0.70), suggesting the two cellular assays are potentially interchangeable. The suboptimal concordances between rat in vivo repeated dose and cancer cell lines were markedly improved by 10%, when limiting the compounds causing drug-induced liver injury (DILI) endpoints and a gene list of DILI predictive toxicogenomics space (PTGS). Furthermore, some toxicity related pathways including PPAR signaling pathways, and fatty acid-related pathways were consistently perturbed across the assay systems, confirming our previous finding that assay transferability is biological process specific. Also, the extrapolation ability for differentially expressed genes (DEGs) of 304 immune-related states among assay systems were evaluated, suggesting the preservation of immune- related transcriptional features is assay dependent. In conclusion, these comparisons suggested a great potential of repurposing transcriptomic profiling assays of immortalized cell lines in DILI.

Keywords: PRank method; Toxicogenomics; CMap; Drug-induced liver injury (DILI)

AR-BIC 22

High throughput kernel phenotyping of corn hybrids grown in Arkansas

Karina Medina-Jimenez[1], Nathan D. Miller[3], Kendl Fischer[1], Zachary C. Campbell[1], Chineche L. Aniemena[1],Cassie Phillips[1], Amber Booth[1], Elizabeth E.Hood[1], Argelia Lorence[1,2]
[1]Arkansas Biosciences Institute, Arkansas State University, Jonesboro, AR;
[2]Department of Chemistry and Physics, P.O. Box 419, State University, AR, 72467, USA
[3]Department of Botany, University of Wisconsin-Madison, 430 Lincoln Drive, Madison, WI 53706, USA

Maize is one of the most important sources of sustenance in many countries around the world. One common way to evaluate the yield of maize crops has been to analyze kernel length, width, and weight; these traits can be strongly affected by growth conditions, especially abiotic stresses such as heat, cold, and drought.To obtain more precise information of the maize kernel's traits, a large number of measurements are necessary. The conventional method for measuring kernel traits is manual, time-consuming, costly, and subjective. During the past few years, high throughput methods have been developed to evaluate the quality of kernels and measure their shape using digital imaging and computer vision. In this study, using a combination of a cyber-infrastructure named CyVerse and an image processing method developed by ND Miller as part of the Phytomorph pipeline we are evaluating geometric parametersof kernels from 250 hybrid corn lines grown in Arkansas as part of the Genomes to Fields (G2F) Project. Those parameters include area, perimeter, kernel tip, long axis, kernel width, and color hues. This process uses images acquired with RGB, near infrared (NIR) and fluorescence (FLUO) sensors that are part of a Scanalyzer HTS instrument. Thousands of these images are being analyzed through segmentation algorithms that are increasing our ability to evaluate in detail kernel phenotypes providing an efficient tool to assist in maize research.

AR-BIC 23

The Sesquiterpene Lactone Parthenin, from Parthenium hysteropherus, Ablates Acute Myelogenous Leukemia (AML)

David E. Mery[1,2,3], Rachel Yang[1,2], Paola E. Ordóñez[3], Eloy Rodriguez[1,], Cesar M. Compadre[3], and Monica L. Guzman[2]
[1]Dept. of Plant Biology, Cornell University, Ithaca, NY 14850, [2]Dept. of Hematology/Oncology, Weill Cornell, New York, NY 10065-4896, and [3]Dept. of Pharmaceutical Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205

Background: Leukemia stem and progenitors cells are central to leukemia relapse by maintaining and initiating the leukemic cell population. The sesquiterpene lactone (SL) parthenolide (PTL) has raised interest as a potential anti-leukemic compound because its ability to target leukemia stem cells. Unfortunately PTL's poor water solubility and relatively low potency limits its in vivo effectiveness. In our continued effort to identify SLs with a better pharmaceutical profile, we have uncovered parthenin (PRT), a SL from Parthenium hysterophorus.
Results: PRT was tested in comparison with PTL against a battery comprised of 12 leukemia cell lines, 4 primary leukemia cell samples, and 3 normal PBMC samples. Compared to PTL, PRT kills AML at lower doses (Mean LD50: 6.81 µM vs. 11.56 µM Cell Lines, and 6.80 µM vs 7.91µM Primary Cells), depletes less free thiols and induces less ROS, inhibits NF-κB transcriptional targets better, causes less activation of Nrf2 transcriptional targets, and decreases active NF-κB and HMOX-1 protein levels better. Mechanistically, it is believed that SLs induce apoptosis through inhibition of NF-κB which is upregulated in leukemia cells. Molecular modeling analysis suggests that both PRT and PTL can bind very well to NF-κB because they have a large lipophilic surface formed by C-8, C-9, C-13 and C- 14 that seem to interact with a complementary lipophilic surface in NF-κB. Tetraneurin-E (TET), a SL that is structurally similar to PRT but has a smaller lipophilic contact surface, was tested against the same battery of cells. Interestingly, TET showed no activity against the leukemia cell lines or the leukemia primary cells that have elevated NF-κB levels, but showed toxicity against the normal cell lines that have normal NF-κB levels.
Conclusions: These results support the hypothesis that the SLS's mechanism involves interaction with NF-κB, and stress the importance of the presence of an extended lipophilic surface for their activity.

AR-BIC 24

Mechanisms Underlying the Enhanced Biomass and Abiotic Stress Tolerance Phenotype of Arabidopsis MIOX Over-expressor Line

Nirman Nepal[1], Jessica P. Yactayo-Chang[1], Karina Medina-Jiménez[1,2], Lucia M. Acosta-Gamboa[1], Mario A. Arteaga-Vazquez[2], and Argelia Lorence[1,3]
[1]Arkansas Biosciences Institute, Arkansas State University, P.O. Box 639, State University, AR 72467, USA
[2]INBIOTECA, Universidad Veracruzana, Avenida de las Culturas Veracruzanas No. 101, Colonia Emiliano Zapata, CP 91090, Xalapa, Veracruz, México
[3]Department of Chemistry and Physics, Arkansas State University, P.O. Box 419, State University, AR 72467, USA.

Myo-inositol oxygenase (MIOX) is the first enzyme in the inositol route to ascorbate (L-ascorbic acid, AsA, vitamin C). We have previously shown that Arabidopsis plants constitutively expressing MIOX have elevated foliar AsA content and displayed enhanced growth rate, biomass accumulation, and increased tolerance to multiple abiotic stresses. In this work we used a combination of transcriptomics, chromatography, microscopy, and physiological measurements to gain a deeper understanding about the underlying mechanisms mediating the phenotype of the AtMIOX4 over-expressor line. In silico and RT-qPCR analysis revealed increased expression of genes involved in auxin synthesis, hydrolysis, transport, and metabolism, which are supported by elevated auxin levels both in-vitro and in-vivo, and by assays demonstrating their effect on epidermal cell elongation in the AtMIOX4 over-expresser. Additionally, we detected upregulation of transcripts involved in photosynthesis and in support to this finding, the transgenic plants displayed increased efficiency of the photosystem II and proton motive force. We also found increased expression of amylase leading to higher intracellular glucose levels. Multiple gene families conferring plants tolerance to cold, water limitation, and heat stresses were found to be elevated in the MIOX4 line. Interestingly, the high AsA plants also displayed upregulation of transcripts and hormones involved in defense including jasmonates, defensin, glucosinolates, and transcription factors that are known to be important for biotic stress tolerance. These results overall indicate that elevated auxin and glucose, enhanced photosynthetic efficiency in combination with up-regulation of abiotic stresses response genes are some of the reasons that explain the higher growth rate and abiotic stresses tolerance phenotype of the AtMIOX4 over-expressers.

AR-BIC 25

Antioxidant Tocols as Potent Radiation Countermeasures

Ujwani Nukala[1,2], Awantika Singh[1,2], Shraddha Thakkar[3], Mahmoud Kiaei[4], Rupak Pathak[1], Nukhet Aykin-Burns[1,5], Philip J. Breen[1,5] and Cesar M. Compadre[1,5]
[1]Department of Pharmaceutical Sciences, College of Pharmacy, University of Arkansas for Medical Sciences, Little Rock, AR 72205,
[2]Joint Bioinformatics Graduate Program, University of Arkansas at Little Rock, Little Rock, AR 72204,
[3]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079,
[4]Department of Pharmacology and Toxicology, University of Arkansas for Medical Sciences, Little Rock, AR, [5]Tocol Pharmaceuticals, LLC, Little Rock, AR 77205, USA

Purpose: The risk of radiation exposure has been increasing with increased use of ionizing radiation for nuclear power plants or nuclear weapons, both of which can result in accidental radiological emergencies. The vitamin E family consists of eight vitamers, including four tocopherols (α, β, γ and δ) and four tocotrienols (α, β, γ and δ) and their anti-oxidant potential may be a major factor for their radioprotective activity. Though alpha-tocopherol was extensively studied in the past, tocotrienols have recently gained attention as radiation countermeasures. Despite several studies performed on tocotrienols, there is no clear evidence on the factors that are responsible for their superior radiation protection properties over tocopherols and their dramatic differences in their pharmacokinetic and pharmacological profile, despite relatively minor structural differences.
Methods: To test our hypothesis that the flexibility in the tail of tocols is crucial for effective binding to ATTP, which is responsible for plasma bioavailability, we performed molecular dynamic simulations of ATTP (PDB ID: 1OIZ) in complex with tocols were performed using GROMACS 5.1.4. Cell uptake levels of α, γ and δ tocopherols and tocotrienols were determined in HUVEC Cells, NSC-34 cells and HepG2 cells. The anti-oxidant potential of the tocols was evaluated by measuring their ability to inhibit lipid peroxidation in microsomes using a TBARS assay.
Results: The results of our analysis show that not all the tocols are created equally and suggest a paradigm in which the observed differences can be explained by a multifactorial function: **"Tocol's therapeutic efficacy = Fn (Intrinsic bioactivity, Elimination rate, Cellular uptake)"**
Conclusions: The inconclusive outcomes of various vitamin E clinical trials may be due to limited understanding of the pharmacokinetics of tocols and of the fact that the bioavailability of vitamin E in humans depends on various factors. This paradigm can be used to identify or develop novel vitamin E analogues that show rapid cell-uptake and enhanced bioavailability while maintaining the strong radioprotective activity of the tocotrienols.

AR-BIC 26

Human Computer Interface Guided Approach to Accelerate Drug Discovery

Meenhaj Kabir and Rajendram Rajnarayanan

New York Institute of Technology at Arkansas State University, Jonesboro, AR

Structure-based drug discovery of target specific drugs greatly rely on the existence of high resolution X-ray crystal structure of the target proteins.
Flexible and dynamic regions including hinges and loops which constitute major protein-protein interaction sites as well as allosteric sites are often beyond the scope of current tools and techniques available for protein folding and modeling. Human ability in recognizing patterns and to solve complex puzzles are far superior to any existing computer program at folding these atypical regions of proteins. We have utilized an unconventional combination of using dynamic three-dimensional protein models as physical human computer interface (HCI) devices and integrated proteomics data to predict flexible and dynamic protein-protein interfaces and allosteric pockets of key regulatory proteins to accelerate compound discovery. To this end, we have successfully utilized 1) Flexible HCI devices to generate an ensemble of dynamic three-dimensional structures which includes a subset of biologically active conformations among others (thereby exploring the viable chemical space) and 2) Structure refinement and efficient filtering of biologically active conformations can be accomplished by integrating protein-protein interface and fold proteomics data. Streamlining of HCIguided tools to enable access to dynamic druggable pockets in protein targets will accelerate drug discovery.

AR-BIC 27

Application of Bayesian Models to Investigate the Genetics of Leaf Drought Tolerance in Cowpea at Early Vegetative Stage

Waltram Second Ravelombola and Ainong Shi
Department of Horticulture, PTSC 316, University of Arkansas, Fayetteville, AR 72701.

Cowpea [*Vigna unguiculata* (L.) Walp.] is a legume crop providing good quality protein to human consumption. In the United States, cowpea cultivation can be found in the southern and western part of the country. Even though cowpea market value is less than other crops such as soybean, its excellent adaptation to poor conditions and small genome size are substantial assets to investigate the genetic mechanisms underlying biotic and abiotic stresses in legumes. Drought has been one of the major factors constraining crop production worldwide and some cowpea genotypes have been proven to be well adapted to drought conditions. Therefore, this study aimed to conduct association mapping (AM) and genomic selection (GS), to identify single nucleotide polymorphism (SNP) markers, and to assess GS accuracy for leaf drought tolerance traits in cowpea. A total of 331 cowpea genotypes with 6 plants per genotype were evaluated under severe drought conditions in the greenhouse at early vegetative stage. More replications are under investigation. Of the 331 cowpea genotypes, 229 genotypes have been genotyped using a total of 1,049 SNPs postulated from genotype-by-sequencing (GBS). Phenotypic data consisted of leaf greenness score, number of plants affected by unifoliate chlorosis, first trifoliate chlorosis, dead plants, dead growing points, and ratio of trifoliate leaf chlorosis to trifoliate green leaves. A Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) model was used to conduct AM and two Bayesian models were used to conduct GS. Results indicated that a large variation was observed in each leaf phenotypic trait among the cowpea genotypes. We expect some SNPs with LOD >2 for all traits and moderate to high GS accuracy upon completing AM and GS. The results will provide insights to the genetics of drought tolerance in cowpea and can be used to enhance cowpea breeding and other legume programs aiming at developing drought-tolerant crops.

AR-BIC 28

Structural Changes Due to Antagonist Binding in Ligand Binding Pocket of Androgen Elucidated by Molecular Docking and Dynamic Simulations

Sugunadevi Sakkiah and Huixiao Hong*
(* Correspondence at Huixiao.Hong@fda.hhs.gov, 870-543-7296)
Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U. S. Food and Drug Administration, Jefferson, AR, United States

Androgen receptor (AR) is a ligand-dependent transcription factor and a member of the nuclear receptor superfamily. It plays a vital role in the male sexual development and regulates genes expression in a variety of tissues. Binding of small molecules to AR initiates the conformational changes in AR-LBD (ligand binding domain) that affect the binding of co-regulator protein and DNA to AR-LBD. The structural details of WT (wide-type)-AR-LBD bound with antagonists would be very useful in the drug discovery research. Unfortunately, WT-AR-LBD complex with an antagonist is not available in the protein data bank. Hence, we have applied an integrative approach that combines the molecular docking and molecular dynamics simulations techniques to identify the important residues involved in the structural changes due to the antagonist binding. Molecular docking was carried out to find a suitable binding orientation of the antagonist (Bicalutamide) in the ligand binding pocket of AR-LBD. The complexes of WT-AR-LBD with Bicalutamide, WT-AR-LBD with agonist (R1881), and mutant-AR-LBD with antagonist (Bicalutamide) obtained from molecular docking were optimized through 1 µs molecular dynamics simulations to identify the conformational changes in the AR-LBP and the activation function 2 (AF2) site. The results revealed that the binding of the antagonist in WT-AR-LBD, moves the residues such as Val716/Lys720/Gln733/ Met734/Gln738/Glu897 which disrupted the positive and negative clumps of the AF2 site. This disruption of the AF2 site is key for understanding the impact of antagonist binding on subsequent coregulator binding. In conclusion, the structural changes elucidated in our study could be helpful to gain a structural insight of WT-AR-LBD-antagonist and help to design the potent antagonist for AR.

AR-BIC 29

User-friendly programs to identify gene variant information for Medullary Cystic Kidney Disease

Kate Sanders[1,2], Qingbo Wang[1,3], Dr. Jamie Marshall[1,]
[1]The Broad Institute of MIT and Harvard, Cambridge, MA
[2]Department of Mathematics and Computer Science, Hendrix College, Conway, AR
[3]Department of Biomedical Informatics, Harvard Medical School, Cambridge, MA

Kidney disease is the 9th leading cause of death in the United States, primarily due to late diagnosis. Determining causal variants for kidney disease could lead to reduced patient mortality through earlier diagnosis and personalized treatments. However, using multiple web browsers to find information about variants is time-consuming and error-prone. We created 2 user-friendly programs to gather, filter, and annotate gene variant information. We also ran these programs for MUC1, a gene associated with Medullary Cystic Kidney Disease.
1) Mendelian Program - After receiving a phenotype as input, this R package (GitHub: SandersKM/MendelianVariants) returns files of data about related diseases and genes. A researcher interested in specific gene variants can download a CSV from gnomAD, then use my program to remove unwanted variants, add variant annotations, and sort variants by the number of pathogenicity scores passed. For MUC1, we began with 899 gnomAD variants and identified 33 variants that passed all 7 scores, indicating a high likelihood of pathogenicity.
2) Complex Disease Program – Joins 3 sources of information: GWAS hits to associate variants with complex traits, eQTL to show a variant's effect on gene expression in kidney tissues, and linkage disequilibrium for fine-mapping. Previously, no web pages visualized all 3 types of kidney disease gene data.
These tools will enable researchers to discover interesting variants more effectively. Currently, the R package for the Mendelian Pipeline is available online, and a package for and Complex Disease Framework is under development.

AR-BIC 30

Directed Genome Evolution for Macrophage-killing in Staphylococcus agnetis an Agent of Bacterial Chondronecrosis with Osteomyelitis in Broilers.

Abdulkarim Shwani, Nnamdi Ekesi, Sura Zaki, Douglas Rhoads
Cell and Molecular Biology program, University of Arkansas.

Staphylococcus agnetis is a coagulase-variable, Gram positive bacterial species which has been previously associated with subclinical or mild clinical cases of mastitis in dairy cattle. Since we first reported the isolation of this staphylococcal species from the bones and blood of lame broilers at the University of Arkansas, others have identified this same species in chickens. We have demonstrated transmission of bacterial chondronecrosis with osteomyelitis (BCO) through aerosols or in drinking water. We have identified particular BCO isolates that can induce very high incidence of lameness. BCO primarily affects the growth plate in the proximal femur and tibia, the fast-growing leg bones. The annotated complete genome of hypervirulent strain 908 has been published. That genome has been compared to nine genomes we assembled for hypovirulent isolates dairy cattle. Phylogenomic analyses of chicken and cattle isolates of S. agnetis and Staphylococcus hyicus suggest a very close relationship between the cattle and chicken isolates. The hypervirulent chicken isolate, 908, clustered with two of the cattle isolates, including strain 1379. A catalogue of gene differences between the cattle and chicken isolates was constructed using reciprocal blast analyses at the nucleotide and polypeptide level. More than 40 genes and 3 plasmids from strain 908 are absent or poorly conserved in any of the cattle S. agnetis isolates. We have found that whereas strain 1379 is efficiently killed by chicken macrophage, strain 908 not only survives phagocytosis by chicken macrophage, it kills the macrophage within 2 days in culture. We have therefore employed Directed Genome Evolution (DGE) to identify the determinants of macrophage survival and killing. DNA from strain 908 was electroporated into strain 1379 which was then passaged through chicken macrophage to select for resistant bacteria culture that kill chicken macrophage. Through multiple rounds we have identified a small 2260 bp plasmid in 908 that appears to be sufficient for survival and killing. This plasmid contains three predicted Open Reading Frames (ORF). One ORF is predicted to be a replication protein, with no function predicted for the other two. Current work is aimed at determining which ORF(s) are required for survival and killing. Future work would be aimed at determining the roles of this polypeptide in cell killing and in virulence in broilers.

AR-BIC 31

Using 16S rRNA for Microbial Classification: A Cautionary Tale

Visanu Wanchai and David W. Ussery
Arkansas Center for Genomic Epidemiology & Medicine and The Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, 72205

For more than four decades, the 16S rRNA gene has been recognized as a “gold standard” marker for evaluating the phylogenetic relationships of prokaryotes. Ever since the first complete bacterial genome has been sequenced in 1995, the molecular revolution has changed the way of study of microbial phylogeny and community. Early genome sequences revealed multiple rRNA operons; currently on average there are about 3 or 4 rRNA operons per genome, but some genomes contain more than 30 copies of the rRNA genes. Although often the 16S rRNA sequences are more than 99% identical to each other in the same genome, in some cases they can vary. The presence of intragenomic heterogeneity of 16S rRNA genes represents a challenge in the estimation of prokaryotic diversity in a complex environment. Here, we present the diversity of multiple 16S rRNA gene copies as well as an analysis of similarity bias between full-length and different regions of 16S rRNA genes. A collection of 10,740 complete and high-quality prokaryotic genomes in the National Center for Biotechnology Information (NCBI) database in January 2019 was downloaded and scanned for rRNA genes, of which 9,329 genomes possessed 2 to 27 copies of 16S rRNA genes per genome. The majority of genomes have near-identical 16S rRNAs (99%); however, we have detected some cases where there are more than one cluster of 16S rRNA genes within a genome with sequence similarity score less than 83%. And in some organisms, different rRNA operons are differentially expressed in different environmental conditions (such as different temperatures). In conclusion, intragenomic heterogeneity of 16S rRNA genes provides evidence supporting the cautious use of 16S rRNA genes to classify the number of prokaryotic species in complex microbiomes. Also, these results might improve our understanding regarding the evolution for each set of copies of 16S rRNA gene within prokaryotic species.

AR-BIC 32

Lifelong dynamics of the swine gut microbiome: from birth to market

Xiaofan Wang[1]; Tsungcheng Tsai[1], Charles V. Maxwell[1], and Jiangchao Zhao[1*].
[1]Department of Animal Science, Division of Agriculture, University of Arkansas, Fayetteville, AR

Despite the recent advances in the swine gut microbiomes during different growth stages, a comprehensive longitudinal study of the lifelong dynamics of the swine gut microbiome is lacking. To fill this gap of knowledge, we selected seventeen piglets (PIC29*380) that were born on the same date from three sows. We collected a total of 273 rectal swabs during lactation (d 0, 11, 20), nursery (d 27, 33, 41, 50, 61), growing (d 76, 90, 104, 116), and finishing (d 130, 146, 159, and 174) stages. Samples were extracted using the Powersoil DNA isolation kit (Qiagen, Hilden, Germany) and sequenced with an Illumina Miseq sequencer targeting the V4 region of the 16 S rRNA gene. Sequences were analyzed with the Deblur algorithm in the QIIME2 package. In general, alpha diversity including community richness (e.g., number of observed features, Chao1) and diversity (e.g., Shannon Index) showed an overall trend of increasing from lactation to the finishing stage. Gradual and significant changes in community structures were also observed along the four growth stages (ANOSIM, $R = 0.66$; $P < 0.01$). Non-parametric permutational multivariate analysis of variance shows that main factors driving the lifelong community dynamics included age and sow origin. Seventeen phylum members were discovered in the lifelong pig gut microbiome with Firmicutes and Bacteroidetes being the most abundant phyla. LEfSe analysis revealed 63 bacterial features that are stage specific. By using a regressing tree based Random Forest model we identified five bacterial features that are associated with swine growth performance including features 26 (Turicibacteraceae Turicibacter), 27 (Clostridium butyricum), 55 (Dialister), 75 (Clostridiaceae) and 4 (Clostridiaceae). Characterization of the lifelong dynamics of 17 healthy pigs from birth to market provides a foundation for gut microbiome studies focusing on swine development, health and growth performance.

AR-BIC 33

Discontinuation of trichloroethylene attenuated the DNA methylation changes in polycomb protein binding sites in effector/memory CD4+ T cells observed after continuous developmental exposure

Stephanie D. Byrum [1, 4], Charity L Washam [1, 4], John D. Patterson[2], Kanan Vyas[3], Kathleen M. Gilbert[4], Sarah J. Blossom[3]
[1]Department of Biochemistry and Molecular Biology; [3]Department of Pediatrics; [4]Department of Microbiology and Immunology, [1, 2, 3, 4]University of Arkansas for Medical Sciences, [3, 4]Arkansas Children's Research Institute, Little Rock, AR, USA

Trichloroethylene (TCE) is an industrial solvent and drinking water pollutant associated with CD4 T cell-mediated hypersensitivity and autoimmunity in humans and rodents. In a previous study, we determined that cessation of TCE exposure during adulthood after a developmental exposure did not prevent immunotoxicity or autoimmune pathology. To determine whether these persistent effects were linked to epigenetic changes we conducted whole genome reduced representation bisulfite sequencing (RRBS) to evaluate methylation of CpG sites in autosomal chromosomes in effector/memory CD4 T cells. The mice were exposed to vehicle control or TCE in the drinking water beginning at gestation until ~37 weeks of age or postnatal day (PND) 259. In a separate group of mice, TCE exposure was discontinued at ~22 weeks of age (PND 154). We identified H3 lysine 27 (H3K27)-associated CpG islands involved in transcriptional repression, suggesting that hypermethylation of these regions may prevent the transcription of important genes involved in silencing gene expression. When CpG sites were overlapped with the Open Regulatory Annotation database (ORegAnno), only continuous TCE treatment resulted in 129 differentially methylated regions (DMRs) including 12 unique transcription factors and regulatory elements. Eighty percent of these DMRs occurred in areas known to bind Polycomb group (PcG) proteins that form the polycomb-repressive complex 2 (PRC2) in mammalian cells, namely, SUZ12, EZH2, JARID2, and MTF2. Pathway analysis of the DMRs indicated that TCE primarily altered the methylation of genes associated with regulation of cellular metabolism and cell signaling pathways. Our results demonstrated that TCE differentially methylated binding sites of PcG proteins in effector/memory CD4+ cells with minimal yet potentially biologically significant effects occurring when exposure was discontinued. These results point toward a novel mechanism by which chronic TCE exposure could modulate CD4 + T cell function and promote autoimmunity.

AR-BIC 34

AR-BIC 35

A pan-cancer study analysis reveals novel regulations of *LRP1B* in cancer pathways

Yifan Zhang[1], Mary Yang[1]
[1*]Dept. of Information Science, University of Arkansas at Little Rock, Little Rock, AR, 72204

Background: *LRP1B* is a protein-coding gene encoding a member of the low-density lipoprotein receptor protein family. This gene is a putative tumor suppressor and down-regulated in several cancer types. Presently, its regulatory roles and involved pathways in the disease remain to be elucidated

Results: We compiled somatic mutation and transcriptome profiles for a pan-cancer dataset that contains diverse tumor samples from more than 30 cancer types and over 10,000 patients. We found that the mutation rate of *LRP1B* was significantly higher than the median of the top 150 mutated genes across 33 cancer types ($p = 0.04$). Out of 24 cancer types that have expression profiles for both tumor and normal tissue samples, our differential analysis suggested that *LRP1B* was significantly under-expressed in nine cancer types while over-expressed in three cancer types (FDR < 0.05). Furthermore, we identified a set of 16 genes that were connected with *LRP1B* within the same co-expression modules in at least four cancer types by cancer-type specific co-expression network analysis. Each of the 16 genes involves in one or more known cancer pathways. Additionally, a number of cancer-related signal pathways and biological processes are significantly enriched of these genes, such as *Wnt* signal pathway (adjusted $p = 0.0023$), *ERBB* signal pathway (adjusted $p = 0.0097$) and regulation of cell death (adjusted $p = 0.0021$). Finally, integrating cancer-type specific modules, we constructed a regulatory interaction network consisting of *LRP1B* and the 16 genes.

Conclusion: Our study reveals putative regulation relationships between *LRP1B* and known cancer-related signal pathway genes in various cancer types. We inferred a pan-cancer network of *LRP1B* with its associated genes, which can help us to better understand the regulatory roles of *LRP1B* in the disease development.

AR-BIC 36

Evaluation and Association Analysis of Downy Mildew Resistance in USDA Spinach Germplasm

Bazgha Zia, Gehendra Bhattarai, Chunda Feng, Wei Zhou, Jun Qin, Maria Isabel Villarroel-Zeballos, Yuejin Weng, Waltram Ravelombola, Jim Correll, Ainong Shi, and Beiquan Mou
University of Arkansas, Fayetteville, AR 72701, USA; and Crop Improvement and Protection Research Unit, USDA-ARS, Salinas, CA 93905, USA.

Downy mildew, caused by the plant fungal pathogen Peronospora farinosa f. sp. spinaciae (Pfs), is an important disease of spinach, particularly for Spinacia oleracea, the most commonly cultivated spinach species in United States and Europe. To date, 16 races of the downy mildew (DM) pathogen are recognized but new isolates are continuously emerging each year. The ability of new strains of the pathogen to overcome resistance in spinach plants makes the development of spinach varieties with increased levels of resistance to Pfs challenging and essential. Genome wide association mapping offers a promising tool to identify QTLs associated with downy mildew resistance against emerging new races of Pfs. The identified QTLs can be readily used to improve genetic resistance against downy mildew disease in spinach. A total of 481 spinach genotypes are evaluated to identify potential resistant germplasm to be used in spinach breeding programs. The evaluated population will be subjected to genome wide association mapping. As a result of this study, SNP markers for downy mildew resistance will provide breeders with a tool to select resistant plants and lines in spinach against downy mildew disease.