4th Annual Meeting of the Arkansas Bioinformatics Consortium
AR-BIC 2018

We are an Arkansas Collaborative Community in Bioinformatics Research

Data Analytics for Genomics and Beyond

Embassy Suites Hotel
Little Rock, AR

April 23-24th, 2018

Organized by:

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

Arkansas Bioinformatics AR BIC Consortium

ArkansasResearchAlliance
Coming together to move Arkansas forward.

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

TriNetX

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 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

Venue:

Ambassador I - IV
Embassy Suites
11301 Financial Centre Pkwy
Little Rock, Arkansas 72211, USA

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.
Staff Scientist , 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
Data Analytics – Genomics and Beyond
April 23-24th 2018

Day 1: Monday April 23rd 2018

12:00 pm – 4:00 pm	**Registration and Poster set up**
2:00 pm – 4:00 pm	**TriNetX Workshop** Mark Hogendobler, Director of Sales, TriNetX, Inc., Austin, TX
4:00 pm – 4:30 pm	**Break**
4:30 pm – 7:30 pm	**Poster Session and Reception**

Day 2: Tuesday April 24th 2018

7:00 am – 8:00 am	**Registration and Breakfast**
8:00 am – 8:15 am	**Welcome and Opening Remarks** • Jerry B. Adams, President/CEO, Arkansas Research Alliance, Conway, AR • Stephanie Gardner, Pharm.D., Ed.D., Interim Chancellor, University of Arkansas for Medical Sciences, Little Rock, AR
8:15 am – 11:45 am	**Session 1: Data Analytics for Genomics** **Session Co-Chairs** • David W. Ussery, Ph.D., Professor, Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR • Fred Prior, Ph.D., Professor and Chair, Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR
8:15 am – 8:45 am	**MG-RAST, a Metagenomics Service for Analysis of Microbial Community Structure and Function** Folker Meyer, Ph.D., Argonne National Laboratory, University of Chicago, Argonne, IL
8:45 am – 9:15 am	**What is Life? Five-hundred Functional Domains Found in All Genomes Across the Tree of Life** David W. Ussery, Ph.D., Professor, Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR
9:15 am – 9:45 am	**Whole Genome Resequencing for Mapping the Genetics of Hypertension in Broilers** Douglas Duane Rhoads, Ph.D., Professor, Department of Biological Sciences, University of Arkansas, Fayetteville, AR
9:45 am – 10:15 am	**Break**
10:15 am – 10:45 am	**Precision Medicine Clinical Trials and the Role of Liquid Biopsies** Donald Johann, Jr, M.D., Associate Professor, Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR
10:45 am – 11:15 am	**Integrative Analysis Identifies Potential DNA Methylation Biomarkers for Pan-Cancer Diagnosis and Prognosis** Tieliu Shi, Ph.D., East China Normal University, Shanghai, China
11:15 am – 11:45 am	**RNA-seq and its Applications in Lung Cancer** Mary Yang, Ph.D., Associate Professor, University of Arkansas at Little Rock, Little Rock, AR
11:45 am – 1:00 pm	**Lunch**

1:00 pm – 5:00 pm	**Session 2 – Data Analytics Beyond Genomics** **Session Co-Chairs** • Weida Tong, Ph.D., Director, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US FDA, Jefferson, AR • John R. Talburt, Ph.D., Professor, University of Arkansas at Little Rock, Little Rock, AR
1:00 pm – 1:30 pm	**Establishing an Enterprise Information Management** Zak Joundi, B.S., Midcontinent Independent System Operator (MISO)
1:30 pm – 2:00 pm	**Disruptive Technology, and Business Model Trends** Gary Dowdy, B.S., MBA, Head of Innovation and Disruptive Technology J. B. Hunt
2:00 pm – 2:30 pm	**Beyond Connecting the Dots: Network Clustering Algorithms and Their Applications** Xiaowei Xu, Ph.D., Professor, University of Arkansas at Little Rock, Little Rock, AR
2:30 pm – 3:00 pm	**Break**
3:00 pm – 3:30 pm	**Real World Text Mining at FDA/NCTR** Joe Meehan, Computer Scientist, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US FDA, Jefferson, AR
3:30 pm – 4:00 pm	**Natural Language Processing – Techniques and Applications to Bioinformatics** Susan E. Gauch, Ph.D., Professor, Department of Biological Sciences, University of Arkansas, Fayetteville, AR
4:00 pm – 4:30 pm	**From *in silico* Drug Design and Discovery to Commercialization** Cesar Compadre, Ph.D., Professor, University of Arkansas for Medical Sciences, Little Rock, AR
4:30 pm – 5:00 pm	**Regulatory Applications of Quantitative Structure Activity Relationship (QSAR)** Weida Tong, Ph.D., Director, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US FDA, Jefferson, AR
5:00 pm – 5:10 pm	**Closing Remarks** Weida Tong, Ph.D., Director, Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US FDA, Jefferson, AR

List of Poster Presenters

Presenter Name	Organization	Title of the Presentation	Poster Number	Abstract Page Number
Jason Abernathy	USDA-ARS	Coordinated effort to advance genomes-to-phenomes through the integration of bioinformatics with aquaculture research	AR-BIC - 1	45
Hoda Hagrass	UAMS	Data Quality of Arkansas clinical data repository (AR-CDR)	AR-BIC - 2	46
Jawaher Alkahtani	UA	Identification and Characterization of Salinity Tolerance Genes by Activation Tagging in Arabidopsis	AR-BIC - 3	47
Stephanie Byrum	UAMS	Characterization of Staphylococcus aureus extracellular protease activity with a protein mass-based assay for proteolysis	AR-BIC - 4	48
Cord Carter	UAMS	Development of an Annotated Dataset that can be used to Analyze and Predict Idiosyncratic Liver Injury	AR-BIC - 5	49
Lauren Corby	UAPB	Study the high-temperature operation of materials for 3D high density power modules	AR-BIC - 6	50
Yheni Dwiningsih	UA	Circadian Expression Patterns of the HYR Gene in Diverse Rice Genotypes	AR-BIC - 7	51
Hong Fang	NCTR	Profiling of Serious Adverse Drug Reactions Using FDA-approved Drug Labeling and MedDRA	AR-BIC - 8	52
Chirag Gupta	UA	Network-based approach to prioritize lung cancer genes from whole exome-sequencing data	AR-BIC - 9	53
Don Johann	UAMS	Scientific and Methodological Advances in Liquid Biopsies to Further the Development of Lung Cancer-Based Precision Medicine	AR-BIC - 10	54
Anuj Kumar	UA	Genome-Wide Association Analyses for Grain Yield and Quality Traits in Japonica Rice under High Nighttime Temperature.	AR-BIC - 11	55
Miles Lange	USDA-ARS	Fish mucus alters the *Flavobacterium columnare* transcriptome	AR-BIC - 12	56
Dongying Li	NCTR	Computational Identification of MicroRNAs That Regulate Sulfotransferase 2A1	AR-BIC - 13	57
Zhichao Liu	NCTR	Transcriptional Responses Reveal Similarities Between Preclinical Rat Liver Testing Systems	AR-BIC - 14	58
Kristin McEuen-Ashby	NCTR-UAMS-UALR	Host Factors and Drug Properties Associated With the Development of Chronic DILI	AR-BIC - 15	59
Darshan Mehta	NCTR	Mining Pharmacogenomic Information from Drug Labeling Using FDALabel Database for Advancing Precision Medicine	AR-BIC - 16	60

David Mery	UAMS	The Sesquiterpene Lactone Parthenin, from Parthenium hysteropherus, Ablates Acute Myelogenous Leukemia (AML).	AR-BIC - 17	61
Yu Nie	UA Little Rock	How Important is the Chief Data Officer for the Company	AR-BIC - 18	62
Joseph Onyilagha	UAPB	Origin of the Genetic Code	AR-BIC - 19	63
Alia Parveen	UA	Whole Genome Resequencing to identify QTLs for ascites in chickens	AR-BIC - 20	64
Abdulkarim Shwani	UA	Directed Genome Evolution to identify Staphylococcal pathogenicity genes for macrophage survival and killing.	AR-BIC - 21	65
Mahanaz Syed	UAMS	Template Based Data Integration and Prediction (TB-DIP)	AR-BIC - 22	66
Visanu Wanchai	UAMS	Analyses of genome quality scores across 120,000 genomes	AR-BIC - 23	67
Dong Wang	NCTR	Infer the in vivo Point of Departure with ToxCast in vitro Assay Data Using a Robust Learning Approach	AR-BIC - 24	68
Leihong Wu	NCTR	Deep learning for food contamination detection	AR-BIC - 25	69
Wei (Vivian) Zhuang	NCTR	A machine learning approach for prediction of the benefit-risk profile with multiple time-to-event endpoints	AR-BIC - 26	70

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. 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 an adjunct appointment at several universities. Also, 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; (4) 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. Also, his group also specializes in molecular modeling and QSARs with a specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book chapters

Shraddha Thakkar, Ph.D.
Staff Scientist
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.

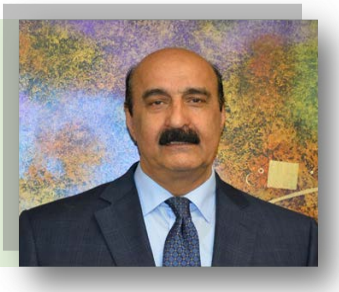
Ralph Davis, Ph.D.
Associate Vice Provost for Research and Innovation
And Professor in the Department of Geosciences
University of Arkansas
Fayetteville, AR

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.

Biographies

Remarks, Co-Chairs, and Organizers

Jerry B. Adams (Welcome Remarks)
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.

Stephanie Gardner, Pharm.D., Ed.D.(Opening Remarks)
Interim Chancellor
University of Arkansas for Medical Sciences
Little Rock, AR

At the University of Arkansas for Medical Sciences (UAMS) Stephanie F. Gardner, Pharm.D., Ed.D., serves as interim chancellor in leading Arkansas' only academic health sciences center, which encompasses patient care, education, research and outreach resources at locations across the state.

Dr. Gardner also currently serves as senior vice chancellor for academic affairs and provost. She has more than 26 years of experience rising through the ranks at UAMS since joining the campus as assistant professor in the Department of Pharmacy Practice in the UAMS College of Pharmacy in 1991. She became chair of that department in 1996, interim dean in the college in 2003 and dean in 2004, leading the College of Pharmacy through unprecedented growth and earning national prominence. She was named associate provost for society and health at UAMS in 2013 and oversaw the vast development of an interprofessional education initiative that became a graduation requirement for all UAMS students. In 2015, she was named senior vice chancellor for academic affairs and provost, providing strategic leadership for the academic and organizational operations of 73 academic programs across the campus's five colleges and graduate school.

A prestigious American Council on Education (ACE) Fellow in 2012-13, Gardner earned Bachelor of Science in Pharmacy and Doctor of Pharmacy degrees from the University of North Carolina, Chapel Hill. She completed a research fellowship in cardiovascular pharmacology from Case Western Reserve University in Cleveland in 1991, and earned a Doctor of Education from the University of Arkansas at Little Rock in 2001.

Gardner's many honors include the Arkansas Pharmacists Association Pharmacist of the Year Award, the Dale Bumpers AHEC (Area Health Education Center) Award and the Outstanding Dean Award by the American Pharmacists Association Academy of Student Pharmacists. Gardner serves as a board member for the Arkansas Center for Healthcare Improvement and the Winthrop Rockefeller Institute. From 2010-16, Gardner was a member of the Accreditation Council for Pharmacy Education Board of Directors and served as president from 2014-15.

David W. Ussery, Ph.D. (Session – 1 Co-Chair)
Professor,
Biomedical Informatics,
University of Arkansas for Medical Sciences,
Little Rock, AR

Professor David Ussery was born and raised in Springdale, Arkansas. He has been working with bioinformatic analysis of bacterial genomes since the first sequence was published in 1995, and published one of the first text books in the field of Comparative Genomics. He has published more than 200 papers, which have been cited more than 10,000 times, including two papers with more than a thousand citations. He has been a co-applicant on grants funded totaling more than $30 million, since 2010. His popular course on Comparative Microbial Genomics, taught at The Technical University of Denmark from 1997 - 2013, is currently running for the 19th year; one-week workshops based on this course have been held in North and South America, Europe, Asia, and Africa. Prof. Ussery has collaborative projects with groups in Belgium, Denmark, France, Germany, The Netherlands, Norway, Spain, Sweden, and the UK, as well as in the U.S.

Prior to joining UAMS, Dr. Ussery was the Comparative Genomics Group leader at Oak Ridge National Labs, in Oak Ridge, Tennessee (2013-2016). He led the Comparative Microbial Genomics group at The Technical University of Denmark from 1997 – 2013, where he has successfully supervised more than 20 Ph.D. students in bioinformatics.

Prof. Ussery received a doctorate in Molecular Biology in1993 from The University of Cincinnati College of Medicine and did a post-doctoral fellowship at Oxford University (1992-1996). He earned his master's degree in biophysical chemistry at the University of New Mexico in Albuquerque. He earned a bachelor's degree in chemistry from William Jewell College (Liberty, Missouri) in 1982, and graduated from Springdale High School (Springdale, Arkansas) in 1978.

Fred Prior, Ph.D. (Session – 1 Co-Chair)
Professor and Chair,
Biomedical Informatics,
University of Arkansas for Medical Sciences,
Little Rock, AR

Fred Prior has been named the inaugural chair of the University of Arkansas for Medical Sciences Department of Biomedical Informatics. Before coming to UAMS he was the Director of the Electronic Radiology Laboratory in the Mallinckrodt Institute of Radiology at Washington University School of Medicine in St. Louis, where he has served as director since 2003. He is also was the director of the Center for High Performance Computing, as well as associate director of the Center for Biomedical Informatics and a research professor of radiology. Prior is the principal investigator for the Cancer Imaging Archive, supported by the National Cancer Institute, which provides researchers, educators and the general public with a vast, freely accessible, open archive of cancer-specific medical images and metadata. Prior holds seven U.S. and international patents and is working with a consortium of investigators on the Human Connectome Project, which is mapping comprehensively the neural pathways of the human brain. Prior received a Master of Science in biomedical engineering at Case Western Reserve University in Cleveland in 1984 and a Ph.D. in computer science at the Illinois Institute of Technology in Chicago in 1992. He served as chief of the Section on Radiologic Computing and Imaging Science at the Pennsylvania State University College of Medicine from 1993 to 1997. Prior spent six years in in medical information management research and development, holding senior management positions at Philips Medical Systems and Eastman Kodak Co., as well as Silicon Valley startups

Weida Tong, Ph.D. (Session – 2 Co-Chair, Closing Remarks, Scientific Liaison)
Director
Division of Bioinformatics and Biostatistics
National Center for Toxicological Research
US FDA
Jefferson, AR

Dr. 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 an adjunct appointment at several universities. Also, 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; (4) 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. Also, his group also specializes in molecular modeling and QSARs with a specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book chapters

John R. Talburt, Ph.D. (Session – 2 Co-Chair)
Professor,
University of Arkansas at Little Rock,
Little Rock, AR

Dr. John R. Talburt is a professor of information science and Acxiom Chair of Information Quality. Talburt is known to not only work continually with students to help them succeed academically, but to also find work environments where they can thrive. Talburt has been instrumental in building the Information Quality Graduate Program at UA Little Rock. Today, the program includes the graduate certificate in Information Quality, the M.S. in Information Quality, and the Ph.D. in Integrated Computing with an Emphasis in Information Quality. Talburt also helped form the Advanced Research in Entity Resolution and Information Quality Laboratory and the new Institute for Chief Data Officers. His contributions and research in the field earned him the Data Management Association International 2008 Academic Award and the MIT Information Quality Program Outstanding Contribution Award in 2006.

Talburt serves as the coordinator for Information Quality Graduate Programs and is the primary advisor for Information Quality graduate students. Past students have noted Talburt's dedication to their academic achievement and willingness to help them find jobs following graduation. Talburt is also a top researcher in his field whose work has placed him at the forefront of Entity Resolution in both corporate and academic settings. Talburt uses this to advantage, by providing students learning and career opportunities. In 2010, Talburt helped to bring the International Conference on Information Quality to UA Little Rock for the first time ever. The conference provided a forum for both researchers and practitioners to exchange knowledge and ideas and served as a learning and networking opportunity for students. The conference was hosted at UA Little Rock again in 2013 and 2017 Talburt also serves as Executive Director of the Entity Resolution & Information Quality Laboratory. The lab engages faculty and graduate students in numerous funded research projects; current external partners include Acxiom Laboratory for Applied Research, Arkansas Science and Technology Authority, and the National Science Foundation.

Shraddha Thakkar, Ph.D. (Scientific Liaison, Poster Session Chair)
Staff Scientist
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).

Julie LaRue (Logistic Liaison)
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

Bryan J. Barnhouse (Logistic Liaison)
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.

Biographies

Speakers (In order of appearance)

Mark Hogendobler, B.S.
Director of Sales,
TriNetX, Inc.,
Austin, TX

Mark Hogendobler received a BA from Princeton University in Molecular Biology, and has worked for the past 20 years in healthcare IT, focusing on the areas of physician/provider documentation, structured reporting, and outcomes analytics. He has served as the Director of Business Development for TriNetX since 2015.

TriNetX Workshop

TriNetX provides an application that interfaces to a clinical data warehouse either directly or via i2b2 and presents users the ability to identify cohorts for future clinical trials and human subjects research. TriNetX also is a network of academic medical centers and pharma companies who collaborate on clinical trials.
https://www.trinetx.com/

Folker Meyer, Ph.D.
Argonne National Laboratory,
University of Chicago
Argonne, IL

Folker Meyer is a Senior Computational Biologist at Argonne National Laboratory and a Professor at the University of Chicago. He is interested in understanding microbial ecology using shotgun metagenomics, and his team builds research software to study microbial communities. Folker is the driving force behind MG-RAST.
His research interests include microbial ecology, distributed high performance computing and big data. He leads the MG-RAST project, providing what is currently the most widely used metagenomics and metatranscriptomics analysis platform. MG-RAST has analyzed over 320,000 data sets, containing 1.2 trillion sequences, with 24,804 registered users, as of February, 2018.
Dr. Meyer is the area lead for microbial community sciences in the Department of Energy (DOE) Systems Biology Knowledgebase project KBase.
In the past he has worked on the GenDB and RAST projects. He is a board member of the Genomics Standards Consortium and a founding member of the Earth Microbiome Project

MG-RAST, a Metagenomics Service for Analysis of Microbial Community Structure and Function

Approaches in molecular biology, particularly those that deal with high-throughput sequencing of entire microbial communities (the field of metagenomics), are rapidly advancing our understanding of the composition and functional content of microbial communities involved in climate change, environmental pollution, human health, biotechnology, etc. Metagenomics provides researchers with the most complete picture of the taxonomic (i.e., what organisms are there) and functional (i.e., what are those organisms doing) composition of natively sampled microbial communities, making it possible to perform investigations that include organisms that were previously intractable to laboratory-controlled culturing; currently, these constitute the vast majority of all microbes on the planet. All organisms contained in environmental samples are sequenced in a culture-independent manner, most often with 16S ribosomal amplicon methods to investigate the taxonomic or whole-genome shotgun-based methods to investigate the functional content of sampled communities. Metagenomics allows researchers to characterize the community composition and functional content of microbial communities, but it cannot show which functional processes are active; however, near parallel developments in transcriptomics promise a dramatic increase in our knowledge in this area as well. Since 2008, MG-RAST (Meyer et al., BMC Bioinformatics 9:386, 2008) has served as a public resource for annotation and analysis of metagenomic sequence data, providing a repository that currently houses more than 150,000 data sets (containing 60+ tera-base-pairs) with more than 23,000 publically available. MG-RAST, or the metagenomics RAST (rapid annotation using subsystems technology) server makes it possible for users to upload raw metagenomic sequence data in (preferably) fastq or fasta format. Assessments of sequence quality, annotation with respect to multiple reference databases, are performed automatically with minimal input from the user (see Subheading 4 at the end of this chapter for more details). Post-annotation analysis and visualization are also possible, directly through the web interface, or with tools like matR (metagenomic analysis tools for R, covered later in this chapter) that utilize the MG-RAST API (http://api.metagenomics.anl.gov/api.html) to easily download data from any stage in the MG-RAST processing pipeline. Over the years, MG-RAST has undergone substantial revisions to keep pace with the dramatic growth in the number, size, and types of sequence data that accompany constantly evolving developments in metagenomics and related -omic sciences (e.g., metatranscriptomics).

David W. Ussery, Ph.D.
Professor,
Biomedical Informatics,
University of Arkansas for Medical Sciences,
Little Rock, AR

Professor David Ussery was born and raised in Springdale, Arkansas. He has been working with bioinformatic analysis of bacterial genomes since the first sequence was published in 1995, and published one of the first text books in the field of Comparative Genomics. He has published more than 200 papers, which have been cited more than 10,000 times, including two papers with more than a thousand citations. He has been a co-applicant on grants funded totaling more than $30 million, since 2010. His popular course on Comparative Microbial Genomics, taught at The Technical University of Denmark from 1997 - 2013, is currently running for the 19th year; one-week workshops based on this course have been held in North and South America, Europe, Asia, and Africa. Prof. Ussery has collaborative projects with groups in Belgium, Denmark, France, Germany, The Netherlands, Norway, Spain, Sweden, and the UK, as well as in the U.S.

Prior to joining UAMS, Dr. Ussery was the Comparative Genomics Group leader at Oak Ridge National Labs, in Oak Ridge, Tennessee (2013-2016). He led the Comparative Microbial Genomics group at The Technical University of Denmark from 1997 – 2013, where he has successfully supervised more than 20 Ph.D. students in bioinformatics.

Prof. Ussery received a doctorate in Molecular Biology in1993 from The University of Cincinnati College of Medicine and did a post-doctoral fellowship at Oxford University (1992-1996). He earned his master's degree in biophysical chemistry at the University of New Mexico in Albuquerque. He earned a bachelor's degree in chemistry from William Jewell College (Liberty, Missouri) in 1982, and graduated from Springdale High School (Springdale, Arkansas) in 1978

Douglas Duane Rhoads, Ph.D.
Professor
Department of Biological Sciences
University of Arkansas
Fayetteville, AR

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, and has served as the Director for the past 12 years. Dr. Rhoads research has produced more 53 journal articles, 8 industry technical reports, and 146 presentations.

Whole genome resequencing for mapping the genetics of hypertension in broilers

Our collaborative consortium has been pursuing the underlying genetics of ascites in broilers. Previously we had used SNP chips and genome wide association studies to pursue loci affecting ascites. These approaches had largely been unsuccessful. Recently, we used whole genome resequencing in our ascites experimental research line. The genome sequence data was used to generate plots of each chromosome for differences in SNP frequencies between resistant and susceptible birds by gender. This approach identified 31 chromosomal regions with potential association with ascites phenotype. The regions in broilers were primarily male associated. Most of the regions contained genes that have been shown to be associated with hypertension or blood physiological parameters in human studies. The first region we have investigated further is associated with the CPQ gene on chromosome 2. This gene produces a plasma carboxypeptidase that may influence the angiotensin-renin pathway. Detailed analysis of this region with a large collection of DNAs demonstrated a strong affect of this region on ascites phenotype in multiple populations. Expression data showed that the mRNA of the CPQ gene was elevated in the birds carrying the resistant allele. The exact role of the CPQ gene in affecting ascites phenotype is not known. The other 30 regions are now under investigation using high throughput SNP genotyping to determine the significance of these additional regions in ascites phenotype. This work will define the most significant loci for breeding against ascites in broiler chickens. Funding for this project was from Agriculture and Food Research Initiative competitive grant number 2015-35203-13380 from the United States Department of Agriculture National Institute of Food and Agriculture.

Donald Johann, Jr, M.D.
Associate Professor of Medicine and Biomedical Informatics
Scientific Director of the UAMS Genomics Sequencing Facility
University of Arkansas for Medical Sciences
Little Rock, AR

Dr. Johann is a physician/scientist, Associate Professor at UAMS and Scientific Director of the UAMS Genomics Sequencing Facility. His scientific focus concerns the application of advanced molecular profiling and high-throughput technologies for the characterization of molecular alterations in cancer cells. Areas of emphasis include next-gen sequencing (NGS), high-resolution identity-based mass spectrometry (proteomics), laser capture microdissection (LCM), bioinformatics, and cancer biology. Previously, he was an assistant investigator at the National Cancer Institute (NCI), Center for Cancer Research (CCR), in the Medical Oncology Branch in Bethesda, MD. Prior to attending medical school he worked as an engineer for the Unisys Corporation for six years, where he directed a team of five engineers on projects involving avionic and systems level (OS, compilers) software design and instrumentation. During this time he also earned a graduate degree in computer science with distinction from Hofstra University. Dr. Johann received his M.D., from Case Western and received a graduate with distinction honors for Computer Applications in Medicine. Following residency he became a postdoctoral research fellow at the NIH/NCI Lab of Pathology, under the mentorship of Dr. Lance Liotta, with a focus on clinical proteomics. He was twice selected for AACR Scholar-in-Training Awards for research work involving novel bioinformatics. Medical Oncology/Hematology fellowships were completed at NIH in the NCI and NHLBI. He has authored ~40 publications and contributed to three patents**.**

Precision Medicine Clinical Trials and the Role of Liquid Biopsies

BACKGROUND: Liquid biopsy methodologies for the purpose of plasma genotyping of cell free DNA (cfDNA) of solid tumors are a new class of novel molecular assays. Such assays are rapidly entering the clinical sphere of research-based monitoring in translational oncology, especially for thoracic malignancies. Potential applications for these blood-based cfDNA assays include: i) initial diagnosis, ii) response to therapy and follow-up, iii) tumor evolution, and iv) minimal residual disease (MRD) evaluation. Advanced molecular diagnostic assays will greatly contribute towards the goals of precision medicine, and especially regarding treatment decisions in the adjuvant setting, where avoiding over-treatment and unnecessary toxicity are paramount.

The use of advanced molecular profiling approaches on individual patient tumor tissues to arrive at rationally-derived therapy decisions for *that* 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 clinical trial protocol has been developed, written and approved by the UAMS IRB, 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, have been 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.

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) clinically relevant animal or tissue models

Tieliu Shi, Ph.D.
East China Normal University
Shanghai, China

Dr. Tieliu Shi received his Master degree in Plant Physiology from Shanghai Institute of Plant Physiology, Academy of Sciences in 1992, Master degree in Computer Science in 1999 and Ph.D degree in molecular biology in 2000 from the University of Louisville, USA. After obtained PhD degree, he pursued bioinformatics research and joined the Bioinformatics Center, Shanghai Institute of Biological Sciences, Chinese Academy of Sciences between 2002 and 2008. He moved to East China Normal University by the end of 2008 and served as full professor. He has various research interests, including 1). Disease gene, disease mechanism and biomarker discovery based on multi-level data integration of omics and clinical information; 2). Methodology developments and applications in the high through-put data (NGS data, proteomic data, etc.); 3). Gene regulatory network prediction and protein-protein interaction network prediction; 4). Drug target, drug efficacy and adverse prediction, including Traditional Chinese Medicine. He has published over 100 peer-review papers.

Integrative Analysis Identifies Potential DNA Methylation Biomarkers for Pan-Cancer Diagnosis and Prognosis

DNA methylation status is generally more stable and occurs earlier than gene expression, thus it could be important signatures for tumor diagnosis, treatment and prognosis. However, the signatures regarding DNA methylation changes for pan-cancer diagnosis and prognosis are less explored. Here we analyzed genome-wide DNA methylation patterns in diverse TCGA cancers with machine learning and identified seven CpG sites that could effectively separate tumor tissue samples from adjacent normal tissue samples in 12 main cancers (1216 samples, AUC > 0.99). Those potential diagnostic biomarkers were further validated in the other 9 different TCGA cancers and 4 independent datasets (AUC > 0.92). Three out of the seven CpG sites are correlated with cell division, DNA replication, cell cycle. We then built a multi-classifier using 12 screened CpG sites, which can effectively distinguish 26 different cancers (7605 samples), and the result was repeated in 6 distinct cancers as well as two disparate tumors with metastases (micro-average AUC > 0.89). Furthermore, a series of potential signatures that could significantly predict the prognosis of tumor patients were identified for each of 7 different cancer via survival analysis (P-value < 1e-4). DNA methylation patterns vary greatly between tumor and adjacent normal tissues, as well as different types of cancers. Our findings demonstrate that tissue specific DNA methylation pattern could be used to predict the primary site of metastatic breast and prostate cancers and the DNA methylation probes identified in this study may aid the decision of clinical diagnosis and prognosis for pan-cancer.

Mary Yang, Ph.D.
Associate Professor
University of Arkansas at Little Rock,
Little Rock, AR

Dr. Mary Yang is a tenured faculty at UALR and Director of the MidSouth Bioinformatics Center and the UALR-UAMS Joint Bioinformatics MS/Ph.D. Program. After completing MSECE, M.S., and a Ph.D. degree supported by a Bilsland Dissertation Fellowship at Purdue University, she joined the National Human Genome Research Institute at the NIH. During her tenure there, she made contributions to large-scale genomics and systems biology research projects, and was Founding Editor-in-Chief of International Journal of Computational Biology and Drug Design, a NIH PubMed indexed journal. She is on the editorial broads of The Journal of Supercomputing and International Journal of Pattern Recognition and Artificial Intelligence. She has published over 50 PubMed-indexed articles and 70 DBLP-indexed computer science papers. Dr. Yang's main research interest is in developing functional genomics and systems biology-based approaches that render a better understanding of the molecular mechanisms underlying complex diseases such as cancer.

Integrative Analysis Identifies Potential DNA Methylation Biomarkers for Pan-Cancer Diagnosis and Prognosis

Lung cancer is one of the most fatal cancers worldwide and non-small cell lung cancer (NSCLC) accounts for over 80% of all lung cancer diagnoses. The 5-year survival rate of early stage lung cancer is significantly higher than that for the metastatic stage (55% versus 4%). However, only 16% of lung cancer cases are diagnosed at an early stage. RNA sequencing (RNAseq) allows high-resolution transcriptome analysis, which can lead to the identification of more effective molecular markers for early diagnosis. We analyzed the RNAseq datasets generated from normal adenocarcinoma in situ and early stage NSCLC tissue samples. Differential expression analysis and machine-learning base feature selection revealed a gene set that can be used for early NSCLC detection. In blind classification for two independent patient cohorts, these signature genes consistently yielded over 98% accuracy for distinguishing early stage lung cancer from normal cases. We further extended our study to RNAseq data for tumor-educated platelets for non-invasive biomarker identification.

Zak Joundi, B.S., MBA
Director of Forward Markets
Midcontinent Independent System Operator (MISO).

Zak Joundi leads the Forward Markets division, responsible for the Financial Transmission Rights (FTR) Market, Capacity Market and Data Services and Analytics at the Midcontinent Independent System Operator (MISO).

Mr. Joundi joined MISO in 2006 in the transmission planning department, followed by roles in Market Administration Engineering where he supported stakeholder processes and day-to-day functions within MISO's Forward Operations division which includes Day Ahead Market Administration, Forward Reliability Assessment & Commitment, and FTR.

Subsequently, Mr. Joundi was heavily involved in congestion management protocols with MISO's neighboring entities as part of his role in the Seams Administration before transitioning to a management role overseeing the administration of the FTR and Pricing function.

Mr. Joundi holds a Bachelor of Science degree in Electrical Engineering from Iowa State University and a Master of Business Administration degree from Indiana University.

Gary Dowdy, B.S., MBA
Head of Innovation and Disruptive Technology
J. B. Hunt

Gary Dowdy is an award-winning technology executive with experience in innovation, e-commerce, digital marketing, software development, systems integration, business development and strategic planning from start-ups to Fortune 20.

As Head of Innovation and Disruptive Technology at J.B. Hunt, Mr. Dowdy is responsible for revenue systems and partnerships to advance the company's digital prowess and transformation to a technology driven provider of innovative supply chain solutions.

Gary has a strong entrepreneurial track record. From 1988 to 1997, he was co-founder and president of three technology companies specializing in the automation of distribution centers and creation of software development products. In 1993, he received the Young Entrepreneur of the Year award, presented by the United States Small Business Administration. From 1998 to 2008, Gary held numerous technology, e-commerce and marketing roles at the executive level at Cardinal Health (Fortune 19) in Chicago, Illinois and Columbus, Ohio. Since 2008, Mr. Dowdy has been a consultant and executive at four startup companies, one of which recently sold for $187 million.

Mr. Dowdy taught an Internet Marketing MBA course at the University of Arkansas, Walton College of Business and at The Ohio State University. He serves on the boards of the Arkansas Academy of Computing and Arkansas School for Mathematics, Sciences and Arts. Gary also served on the Governor Hutchinson's Task Force for the Computer Science Initiative and is a member of the Arkansas Executive Forum.

He received a Bachelor's of Science in Computer Science from the University of Arkansas and an MBA from the Krannert School of Business at Purdue University.

Xiaowei Xu, Ph.D.
Professor,
University of Arkansas at Little Rock,
Little Rock, AR

Xiaowei Xu, a professor of Information Science at the University of Arkansas, Little Rock (UALR), received his Ph.D. degree in Computer Science at the University of Munich in 1998. Before his appointment in UALR, he was a senior research scientist in Siemens, Munich, Germany. His research spans data mining, machine learning, bioinformatics, database management systems and high-performance computing. He has published over 100 papers in peer-reviewed journals and conference proceedings. With over 18,114 citations, he is one of the most cited researchers according to Google Scholar. His groundbreaking work on the density-based clustering algorithm DBSCAN has been widely used in textbooks and software implementations; and has received over 11,692 citations based on Google scholar to date. He is a recipient of the prestigious ACM SIGKDD Test of Time award for his contribution to the density-based clustering algorithm DBSCAN. Recently Dr. Xu has been recognized as a Most Influential Scholar in the field of Data Mining for his "outstanding and vibrant contributions to the field of Knowledge Discovery and Data Mining" by AMiner (https://aminer.org/mostinfluentialscholar/datamining), a free online service for academic social network analysis and mining**.**

Beyond Connecting the Dots: Network Clustering Algorithms and Their Applications

Big data is highly complex with many interrelated elements, which can be modeled as a network by connecting the dots (i.e. the interrelated elements). The network provides us a big picture of the underlying complexity, but it is often too big to discover any useful hidden patterns from millions of nodes and links. This talk first presents the clustering algorithms for large networks, which can efficiently detect different hidden patterns including community structures (which refer to the occurrence of groups of nodes in a network that are more densely connected internally than with the rest of the network), and nodes playing special roles such as hubs that connect different community structures, as well as different types of relationships. Another focus of this talk is about the application of the network clustering algorithms to different networks including social networks, biological networks, as well as to entity co-occurrence network for text mining.

Joe Meehan
Computer Scientist,
Division of Bioinformatics and Biostatistics,
National Center for Toxicological Research,
US FDA,
Jefferson, AR

Joe Meehan, Computer Scientist, FDA/NCTR Joe Meehan serves as a Senior Advisor in the Division of Bioinformatics and Biostatistics at the FDA National Center for Toxicological Research (NCTR). Mr. Meehan began his career as computer programmer at Baptist Memorial Hospital in Memphis before coming to the FDA National Center for Toxicological Research in 1983. He was the Systems and Networks manager from 2001 to 2008, and Software Development manager from 2008 to 2012. He is currently the High Performance Computing coordinator at NCTR, and has spent the last several years coordinating collaborative biomedical informatics software development efforts involving NCTR and other FDA centers. Joe represents NCTR on the FDA High Performance Computing Governance Advisory Board (HPC GAB) and the FDA Chief Information Officer's Information Security Subcommittee (INFOSEC).

Real World Text Mining at FDA/NCTR

Much of the information submitted to the FDA and generated by FDA reviewers remains in relatively unstructured text documents. Researchers at the FDA National Center for Toxicological Research (NCTR) have collaborated with staff at the FDA Center for Drug Evaluation and Research (CDER) to develop tools and techniques to extract key data from regulatory documents so that it can be used to monitor and improve the regulatory process. This presentation will describe several recent and ongoing projects in this area.

Susan E. Gauch, Ph.D.
Professor
Department of Biological Sciences
University of Arkansas
Fayetteville, AR

Dr. Susan Gauch is currently a professor in the Computer Science & Computer Engineering department of the University of Arkansas. She received her Ph.D. from University of North Carolina at Chapel Hill in 1990. While at the University of Kansas, from 1993-2007, her research developed one of the first meta-search engines, ProFusion, the first digital video search engine, and some of the earliest work on personalized search. Four projects resulted in software licenses to industry from the University of Kansas.

Dr. Gauch joined the University of Arkansas in 2007 to become the head of Computer Science and Computer Engineering. She held this position until 2015 after which she rejoined the faculty as a professor. Her current research investigates conceptual and personalized search of academic literature, social network analysis of literature, and text mining approaches to sentiment analysis

Natural Language Processing – Techniques and Applications to Bioinformatics

Unstructured, free text can be a rich source of information for many bioinformatics applications. However, because of the ambiguity present in natural language, it can be difficult to extract knowledge from these types of sources. This presentation will summarize some of the challenges, approaches, and successes of natural language processing (NLP), including classification, text extraction, search, and sentiment analysis.

After a general introduction to the field, two specific projects will be discussed in more detail: semi-automated construction of an amphibian morphology ontology and text extraction from pathology reports. The presentation will conclude with an introduction to some publically-available NLP tools.

Cesar Compadre, Ph.D.
Professor
University of Arkansas for Medical Sciences
Little Rock, AR

Dr. Compadre is a professor at the Department of Pharmaceutical Sciences, of the University of Arkansas for Medical Sciences. He has extensive research experience on the development of bioactive compounds based on naturally occurring compounds, and on the use of molecular modeling in drug design and structure-activity studies. He has published over 90 papers and co-authored more than 70 patents related to the development of bioactive compounds.
He is also the developer of one FDA approved antimicrobial technology, which is commercially used, and he is also co-founder of Tocol Pharmaceuticals, a company focused on the development of enhanced vitamin-E analogues. Dr. Compadre has extensive International research collaborations in Drug Discovery, Global Health and Phytopharmaceuticals. Dr. Compadre has a BSPharm degree, and obtained his Ph.D. degree in medicinal chemistry and pharmacognosy, from the University of Illinois at Chicago. He conducted postdoctoral research on structure-activity relationships studies using molecular modeling at the University of Illinois working with Dr. John M. Pezzuto and at Pomona College working with Professor Corwin Hansch. Additionally, he had a sabbatical experience at NASA Ames Research Center in computer modeling.

From *in silico* Drug Design and Discovery to Commercialization

There is an unmet need for radioprotectors, compounds that protect against radiation injury in the event of radiation accidents or terrorism scenarios. In this context, Vitamin E is a very well-known anti-oxidant that scavenges the free radicals produced by radiation exposure. Vitamin E family includes eight different isoforms including four tocopherols (α, β, γ and δ) and four tocotrienols (α, β, γ and δ), which are collectively known as tocols. The standard vitamin E containing preparation sold in the market is α -tocopherol. The main reason for this is that AT has the slowest rate of elimination (t1/2= 18 h) and thus it can be used for once-a-day administration. However, the therapeutic efficacy of AT evaluated on multiple indications has been disappointing and rather poor. On the other hand, there is a rapidly increasing number of studies that show that the tocotrienols have a much superior biological activity compared to the tocopherols for many indications, including radioprotection. Using in silico techniques, a team of researchers at UAMS were able to design a new kind of vitamin E analogues, named the tocoflexols, which have the both the superior bioavailability of the tocopherols and the superior bioactivity of the tocotrienols. Patent protection was applied and received for the tocoflexols, and an Arkansas company, Tocol Pharmaceuticals, was formed to commercialize tis technology. Tocol Pharmaceuticals is part of the UAMS, business incubator, Bioventures.

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

Dr. 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 an adjunct appointment at several universities. Also, 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; (4) 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. Also, his group also specializes in molecular modeling and QSARs with a specific interest in estrogen, androgen, and endocrine disruptor. Dr. Tong has published more than 250 papers and book

Regulatory Applications of Quantitative Structure Activity Relationship (QSAR)

Considerable scientific, regulatory and popular press attention has been devoted to the Endocrine Disrupting Chemicals (EDCs). A larger number of potential estrogenic EDCs are associated with products regulated by the Food and Drug Administration (FDA), including plastics used in food packaging, phytoestrogens, food additives, pharmaceuticals, cosmetics, etc. Given the huge number of chemicals, many commercially important, and the expense of testing, SAR/QSAR has been considered to be an important priority setting strategy for subsequent experimentation. At the U.S. FDA's National Center for Toxicological Research (NCTR), we have conducted the Endocrine Disruptor Knowledge Base (EDKB) project, of which SAR/QSARs is a major component. We have developed predictive models for estrogen and androgen receptor binding. The strengths and weaknesses of various QSAR methods were assessed to select those most appropriate for regulatory priority setting. This presentation, rather than presenting the work and results of the EDKB program in an exhaustive manner, selectively discusses salient concepts, issues, and challenges, endeavoring to achieve a tutorial outcome. In particular, concepts such as designing training sets, living models, use of QSARs in a regulatory context, predictive model validation, QSAR applicability domain and prediction confidence estimates are among topics to highlight. The concepts are presented and discussed using EDKB program results to provide qualitative and quantitative illustrations and examples. We believe the experience and lessons learned in the EDKB program will prove valuable to practitioners of QSAR should they endeavor to extend predictive systems to real-world regulatory implementations.

Poster Presentation Abstracts

AR-BIC – 1

Coordinated effort to advance genomes-to-phenomes through the integration of bioinformatics with aquaculture research

Jason Abernathy, S. Adam Fuller, Bartholomew Green, Miles Lange, Steve Rawles, Dave Straus and Carl Webster

USDA, Agricultural Research Service, Harry K. Dupree Stuttgart National Aquaculture Research Center, 2955 Highway 130 East, Stuttgart, Arkansas, 72160 USA

Aquaculture is the fastest growing food production system in the world. The research program at the USDA-ARS-SNARC strives to improve the efficiency and sustainability of warmwater U.S. aquaculture. SNARC scientists have impacted the catfish (#1 U.S. aquaculture industry), tilapia (#3) and hybrid striped bass (#4) industries among others throughout decades of research. Our current efforts focus on the hybrid striped bass industry, a major industry for the state of Arkansas. By integrating computational approaches coupled with basic and applied aquaculture research we strive to further advance the hybrid striped bass industry through the genetic improvement of the hybrid parental species, the white bass and striped bass. Toward this effort, our work addresses selective breeding of white bass to improve agriculturally-important traits including growth and alternative feeds utilization. Molecular genetics platforms and bioinformatics pipelines are being developed concurrently. The white bass and striped bass genomes are being sequenced and assembled. Transcriptomic resources across white and hybrid striped bass are being characterized in response to variables including diet formulation, diseases and their treatments, and the aquatic ecosystem. We are mapping microbiome and mycobiome changes throughout the production cycle. With reciprocal crossbreeding we are examining quantitative trait loci, the transmissibility of traits to their hybrids as well as maternal and paternal effects. We also use a dynamic systems approach to aid our understanding of host-pathogen interactions. Our long-term goal is to increase genetic gains via the integration of molecular genetics with our selective breeding program. An update on our progress toward mapping functional variation across our multidisciplinary projects as well as the computational approaches utilized at SNARC will be presented.

AR-BIC – 2

Data Quality of Arkansas clinical data repository (AR-CDR)

Saly Abouelenein, **Hoda A. Hagrass**, Dina M. Elsayed , Ahmad Baghal

The Department of Biomedical Informatics, University of Arkansas for Medical Sciences

Introduction: The results of clinical studies depend on the quality of the collected data (1). Arkansas clinical data repository (AR-CDR) is a rich clinical data source (2) .The evaluation of data warehouse quality is crucial to ensure the accuracy of clinical research results (2)

Methods: The current study is a pilot study aimed to evaluate the of data quality in (AR-CDR). We randomly selected the first one hundred diabetic patients either type 1 (ICD 10; E10) or type 2(ICD10; E11) from UAMS data warehouse. We focused on three data quality dimensions: Completeness, accuracy and validity. We used SAS 9.4 to measure these dimensions.

Results: There was great discrepancy between data warehouse and Epic data in Antidiabetic medication due to using different coding systems so we excluded antidiabetic medications.100 % of the data was valid. Nearly 98.8% of the data was complete. There was 100% concordance in all data elements.

Discussion: Our study is the first study that measures different data quality elements in Arkansas and it is one of the few studies measure more than two elements of data quality.

The results of our study should be interpreted in the context of the following limitations:

1. In this pilot study we selected small sample size and one disease.
2. We encountered a specific constrain when we tried to compare drugs list extracted from data warehouse and the corresponding drugs in EPIC because each database has its own coding systems.

Conclusion: This study revealed that there is a complete concordance between data stored in AR-CDR and Epic data. The data in AR-CDR are complete and valid

Based on the results of this pilot study we confirm that the data warehouse has a high quality for clinical research.

References

1) Dedić, N. and Stanier C., 2016., "An Evaluation of the Challenges of Multilingualism in Data Warehouse Development" in 18th International Conference on Enterprise Information Systems - ICEIS 2016, p. 196.

2) Exploring Data Warehouses and Data Quality". Spotlessdata.com. Retrieved 2017-04-30

AR-BIC – 3

Identification and Characterization of Salinity Tolerance Genes by Activation Tagging in Arabidopsis.

Jawaher Alkahtani and Andy Pereira

University of Arkansas, Fayetteville, AR

Salinity often affects irrigated areas in arid and semi-arid regions of the world. The existence and accumulation of soluble salts in the soil layers limit the growth of crops essential for our food. Salt stress dramatically affects plant growth, plant development, as well as crop yield. *Arabidopsis thaliana* is the plant model that provides a comprehensive knowledge of plant development, genetics and physiology, and response to abiotic stresses such as salinity. The redundancy of genes due to duplication, even in the simple model genome of Arabidopsis, limits the value of knockout (KO) mutagenesis to provide complete information on gene function. 'Gain-of-function' mutants are an alternative genetic tool to identify gene functions for redundant genes, and those with small effect or that respond to an environmental condition. Transposon-mediated 'activation tagging' is an efficient genetic tool that can randomly generate 'gain-of-function' mutants for a large number of genes. In the method used here, the transposable element Enhancer-Inhibitor (En-I/dSpm) system of maize was modified to develop an activation tag (AT) mutant library in Arabidopsis. The mobile I-AT transposon contains a transcriptional enhancer, from the cauliflower mosaic virus (CaMV) 35S promoter, located close to the right border of the transposon. This I-AT element was mobilized to randomly insert into the plant genome by transposition from the T-DNA, and can give rise to mutants differing in the level of overexpression of the adjacent genes. Consequently, the gain-of-function dominant phenotypes generated are displayed by the I-AT plants due to enhanced expression of the gene(s) adjacent to the 35S enhancer. In this study, the I-AT library was used to screen for salt tolerance, identified by enhanced growth or biomass of the tagged mutants compared to the wild-type grown in saline conditions. A number of tagged salt tolerance candidate genes were identified flanking the I-AT insertion, and their tagged genes characterized for their role in salt tolerance**.**

AR-BIC – 4

Characterization of *Staphylococcus aureus* extracellular protease activity with a protein mass-based assay for proteolysis

Stephanie D. Byrum[1], Allister Loughran[2], Karen Beenken[2], Daniel Meeker[2], Lisa Orr[1], Samuel G. Mackintosh[1], Ricky D. Edmondson[1], Alan Tackett[1] and Mark Smeltzer[2]

[1]Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, AR, 72205
[2]Department of Microbiology and Immunology, University of Arkansas for Medical Sciences, Little Rock, AR, 72205

We developed a data analysis method to determine protein degradation using a GelCMS technique and spectral counts. We used a murine model of acute, post-traumatic osteomyelitis to evaluate the virulence of *Staphylococcus aureus* clinical isolates LAC (USA300) and the mutant derivative *sarA* to determine what proteins were being degraded in the *sarA* mutant compared to the parent strain.

The assessment of *S. aureus* secreted proteome of the parent strain LAC and the isogenic *sarA* mutant was analyzed using a Gel-shift spectral count method. Briefly, SDS-PAGE lanes were divided into 24 slices and subjected to in-gel trypsin digestion. Tryptic peptides were analyzed by high resolution tandem mass spectrometry with a Thermo LTQ Orbitrap Velos mass spectrometer coupled to a Waters nanoACQUITY LC system. Proteins were identified by searching the UniprotKB USA300 (LAC) database against *Staphylococcus aureus* (2607 entries).

A total of 1334 proteins were identified from triplicate samples of the LAC and *sarA* mutant strains. Spectral counts were analyzed for each gel slice (1-24) for all six samples. The gel band with the highest spectral count for a given protein in LAC was identified and compared to the corresponding band in the *sarA* mutant. A student's t-test was performed to identify proteins that are significantly different in the mutant compared to the wild-type strain. Proteomic analysis confirmed reduced accumulation of multiple extracellular proteins in LAC and *sarA* mutant, including the alpha class of phenol-soluble modulins (PSMs), which are important determinants of osteoblast cytotoxicity and bone destruction and repair processes in osteomyelitis.

AR-BIC – 5

Development of an Annotated Dataset that can be used to Analyze and Predict Idiosyncratic Liver Injury

Cord Carter[1], Cesar M. Compadre, Ph.D.[1], Weida Tong Ph.D.[2], Shraddha Thakkar Ph.D.[1, 2]

[1]Department of Pharmaceutical Sciences, University of Arkansas for Medical Sciences, Little Rock, AR 72205

[2]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR 72079

Drug Induced liver injury is the leading cause of hepatic failure and one of the major reasons for drugs to be withdrawn from the market. Although, there are multiple predictive models that can trace liver injury to drug dose and solubility, predicting idiosyncratic liver injury remains an unsolved issue. Idiosyncratic liver injury, which is considered to have low occurrence, could produce severe damages particularly when patients are re-challenged. Idiosyncratic liver does not show strong correlation with dose dependent liver injury but show exposure dependence and remains hard to predict at the preclinical stage and is often associated with the human leukocyte antigen (HLA) complex. There have been multiple reports indicating correlation of idiosyncratic liver injury with reactive metabolite formation that can be assessed by their ability to form glutathione adducts and bind to hepatic proteins, produce CYP3A4 inhibition, cause mitochondrial toxicity and cytotoxicity. In this project, we have developed an annotated database that contains optimized 3D structures for over 900 molecules, with over 50 descriptors selected to describe the geometry, reactivity, and physicochemical properties that they may be relevant to liver toxicity. The database can be utilized to develop, analyze, and validate computational models that predict idiosyncratic liver injury.

AR-BIC – 6

Study the high-temperature operation of materials for 3D high density power modules

Lauren Corby, Dr. Fisher Yu and. Dr. Sattar Al-Kabi

University of Arkansas at Pine Bluff

The conventional power modules require a comprehensive thermal management system, which adds weight and cost. High-temperature high-density power modules will not only enhance reliability but also substantially reduce cooling requirements. To enable the entire power modules to operate in the harsh environment, the high temperature operating capability is not only needed for power devices, but also for gate driver circuits, which includes a galvanic isolation system. The typical operating temperature for optocouplers is up to 100^{o}C, due to the limitations of LED devices inside and packaging materials of the optocouplers.

One of the main reasons the operating temperature of optocouplers is limited to a little over 100oC is because of the degradation of LEDs at high temperatures. The degradation mechanism of the LEDs can be identified by studying the electrical and optical characteristics of the LED material at high temperature.

An advanced Photoluminescence (PL) & Electroluminescence (EL) measurement system is built to characterize LED materials at high temperature (~800K). Initial measurements from the PL system provides the optical behavior of LED materials at high temperatures. These results also give an opportunity to study about the efficiency drop at high temperatures in solid state lighting. Excitation dependence PL study using different lasers will provide in-depth knowledge about the nonradiative lifetime and Internal Quantum Efficiency (IQE) of the materials that are under consideration for the high-temperature optocoupler. The objective of the EL study is to characterize LED devices at high temperature. Temperature-dependent study on LED devices will give insight to optical degradation as well as the different failure mechanisms associated with the device.

Using PL/EL measurement system enables the research team to characterize and study the PL for materials and EL for LED devices at high temperature

- Measure PL/EL from 77K-800K
- Excitation power-dependent PL study can be done with two different type of lasers
- A 395nm laser with maximum power 120mW and a (532nm) laser with 500mW are used for PL/EL measurements
- Current-voltage (I-V) characterization will be done for LEDs devices to study the performance at different temperatures.

AR-BIC – 7

Circadian Expression Patterns of the HYR Gene in Diverse Rice Genotypes

Yheni Dwiningsih[1], Julie Thomas[1], Sara Yingling1 and Andy Pereira[1]

1 Crop, Soil and Environmental Sciences Department, University of Arkansas, Fayetteville, AR, USA

The HYR (Higher Yield Rice) transcription factor gene regulates photosynthetic carbon metabolism (PCM) in rice (1). Overexpression of HYR (Os03g02650) in rice lines enhances photosynthetic capacity, and contributes to higher grain yield under normal conditions as well as under drought and high temperature stress. Since HYR is involved in regulation of PCM and response to environmental stress signals we were interested in studying the genetic basis of this response. To study the intrinsic role of HYR in regulating PCM in diverse rice genotypes, the expression profile of HYR was characterized in diverse genotypes at different times of the day. Rice genotypes belonging to the Oryza sativa s.sp japonica (Nipponbare, Kaybonnet, Bengal), and O. sativa s.sp indica (Vandana, Nagina, Aochiu) and a stress tolerant African rice O. glaberrima accession were used for this analysis. Circadian rhythm is an endogenous entrainable oscillation or cycle of around 24 hours, synchronized to the light/dark diurnal cycles that regulates the expression of multiple genes (2), and closely relevant to photosynthetic activities (1). The set of diverse rice genotypes were grown in 500-ml pots in growth chambers (14h light/10h dark cycle with light intensity 580 µmol m-2s-1, with light period from 7am-9pm at 25oC and 65% RH) under well-watered conditions. At the vegetative stage, 48-day-old plants were used for sampling the youngest fully expanded leaves that were collected at 12am, 3am, 7am, 9am, 12pm, 3pm, 6pm and 9pm. RNA was isolated and used to generate cDNA. qRT-PCR experiments were carried out using a CFX-96 Bio-Rad thermocycler. The comparative threshold cycle (Ct) method of quantitation with ubiquitin as reference were used for qRT-PCR analysis. The results show that the expression of HYR was strongly affected by circadian rhythm. The variation of the HYR expression among rice genotypes indicate HYR has different internal time in each rice genotype.

AR-BIC – 8

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

Hong Fang[1], Leihong Wu[2], Taylor Ingle[1], Zhichao Liu[2], Shraddha Thakkar[2], Anna Zhao-Wong[3], Steve Harris[2], Guangxu Zhou[2], Junshuang Yang[1], Joshua Xu[2], Weida Tong[2]

1Office of Scientific Coordination, NCTR
2Division of Bioinformatics and Biostatistics, NCTR
3MedDRA MSSO, McLean, VA

Abstract

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 – 9

Network-based approach to prioritize lung cancer genes from whole exome-sequencing data

Chirag Gupta[1] and Andy Pereira[1]

[1]Department of Crop Soil and Environment Sciences, University of Arkansas, Fayetteville, AR, USA

Sequencing tumor genomes reveal extensive information about disease-causing mutations. However, determining whether the identified anomalies are true driver mutations, or they simply represent passenger mutations that have inconsequential effect on the disease, is still a challenging and non-trivial task. Several lines of evidence indicate defects in pathways and regulatory circuits that drive a cell towards the state of malignancy. Computational analysis ranks mutated genes by the frequency of their occurrence in patient cohorts, often overlooking low-frequency mutations that might play a role in the 'rewiring' of important disease-related pathways. Hence, network analysis methods have garnered interest for illuminating pathways dysregulated in tumor states. We evaluated four major variant callers for their ability to detect and rank known cancer-related mutations, and then used a tissue-aware network-based approach to re-prioritize mutated genes in 28 matched tumor-normal exome-sequencing lung cancer samples. This method takes the frequency of mutation occurrence of neighboring genes in the network into account. Evaluation of this approach against a reference benchmark shows that the accuracy of mutation prediction can be significantly improved by such data integration techniques and can lead to many driver genes and combinations that remain undetected otherwise.

AR-BIC – 10

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

Donald J. Johann, Jr.[1], Mary Yang[2], Karl Walker[3], Andy Pereira[4], Huixiao Hong[5], Erich Peterson[1], Jason Liem[1], Meei Liu[1], Joshua Xu[5], Wenming Xiao[5], Weida Tong[5]

[1]University of Arkansas for Medical Sciences, Little Rock, AR, [2]University of Arkansas at Little Rock, Little Rock, AR, [3]University 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: Liquid biopsy methodologies for the purpose of plasma genotyping of cell free DNA (cfDNA) of solid tumors are a new class of novel molecular assays. Such assays are rapidly entering the clinical sphere of research-based monitoring in translational oncology, especially for thoracic malignancies. Potential applications for these blood-based cfDNA assays include: i) initial diagnosis, ii) response to therapy and follow-up, iii) tumor evolution, and iv) minimal residual disease (MRD) evaluation. Advanced molecular diagnostic assays will greatly contribute towards the goals of precision medicine, and especially regarding treatment decisions in the adjuvant setting, where avoiding over-treatment and unnecessary toxicity are paramount.

The use of advanced molecular profiling approaches on individual patient tumor tissues to arrive at rationally-derived therapy decisions for that 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.

AR-BIC – 11

Genome-Wide Association Analyses for Grain Yield and Quality Traits in Japonica Rice under High Nighttime Temperature.

Anuj Kumar, Sara Yingling, Julie Thomas, Charles Ruiz, Yheni Dwiningsih, Andy Pereira

Department of Crop, Soil, & Environmental Sciences, University of Arkansas, Fayetteville, AR 72701, USA , axk018@uark.edu

Rice (Oryza sativa L.) is an important cereal crop that feeds half the world's population. To meet this growing demand, rice production and its sustainability need to be improved. However, the frequency and intensity of high night temperature, is emerging as a potential threat to the stability of rice production. High temperature affects rice grain yield and quality by: i) high maximum temperature with high humidity that can cause spikelet sterility and reduced grain quality, and (ii) high night temperature (HNT) that reduces assimilate accumulation. To genetically dissect heat tolerance under HNT for grain yield and quality, a panel of 81 japonica rice genotypes of the USDA rice mini core collection (URMC) was screened for the number of filled grains (NOFG) and % chalkiness at the R5 stage (just after anthesis) under HNT (280C) in the greenhouse conditions. We performed genome-wide association (GWA) analyses based on genotyping 200,460 SNP variants across the panel of japonica rice genotypes of the URMC to identify tolerant genotypes and favorable loci for grain yield and quality traits using FarmCPU model. In the GWA analyses, we identified 18 significantly associated SNPs for NOFG and 41 SNPs for % chalkiness under HNT. The results show that grain yield and quality traits within the natural variation available under HNT will help to uncover diverse sources for the traits and characterize genetic loci for rice genetics and molecular breeding research. Improving productivity of the rice crop under HNT remains as one the most important current and future challenges in plant genomics, which can only be helped by a concerted analysis of genes and pathways for grain yield and quality both within the plant and genetic variation for favorable alleles present in nature.

Key words: Japonica rice, High nighttime temperature, Genome-wide association analysis

AR-BIC – 12

Fish mucus alters the Flavobacterium columnare transcriptome

Miles D. Lange, Bradley D. Farmer and Jason Abernathy

Harry K. Dupree Stuttgart National Aquaculture Research Center
USDA-Agricultural Research Service, 2955 Hwy. 130 East Stuttgart, AR, 72160, USA

Columnaris disease which is caused by *Flavobacterium columnare* severely impacts the production of freshwater finfish species. Due to the impact on the aquaculture industry, research efforts to better understand the biological processes of *F. columnare* including the formation of biofilms and their contribution to disease are ongoing. The current work sought to evaluate the planktonic and biofilm transcriptomes of *F. columnare* isolate 94-081 under different growth conditions. Our data shows that fish mucus enhances *in vitro* biofilm formation. Global analysis of *F. columnare* transcriptomes stimulated with or without fish mucus revealed significant variability among the differentially expressed genes (DEGs) of the planktonic and biofilm states. DEGs that were common among all biofilms were enriched for bacterial gene ontology groups such as signal transduction, ligand binding and cellular homeostasis and are likely necessary for biofilm formation. In addition different iron acquisition machinery including TonB dependent receptor and ferroxidase genes were expressed among all biofilms. The increased expression of TonB dependent receptor genes and the identification of siderophore synthesis genes among only mucus-stimulated biofilms help to validate a role for the TonB system as a virulence factor. Other DEGs specific to mucus-stimulated biofilms revealed gene ontology groups associated with ribosome biogenesis and protein translation. The current analysis of *F. columnare* transcriptomes adds valuable information about the basic biological processes that occur during the planktonic to biofilm transition. This work will ultimately allow for a better understanding of how biofilms affect *F. columnare* virulence and initiate disease

AR-BIC – 13

Computational Identification of MicroRNAs That Regulate Sulfotransferase 2A1

Dongying Li, Bridgett Knox, Leihong Wu, Gokhan Yavas, Wenming Xiao, Weida Tong, and Baitang Ning

Division of Bioinformatics and Biostatistics, National Center for Toxicological Research,
US Food and Drug Administration, Jefferson, AR, 72079

Sulfotransferase 2A1 (SULT2A1) is a Phase II drug metabolizing enzyme that plays a critical role in sulfate conjugation of many neurotransmitters, steroids and xenobiotics. SULT2A1-dependent sulfation often reduces the toxicity of substrates by increasing their water solubility and enhancing their elimination. Down-regulation of SULT2A1 is associated with several liver diseases, including cholestasis and primary sclerosing cholangitis. While transcriptional regulation of SULT2A1 expression by nuclear receptors is well-established, it remains elusive how SULT2A1 expression is modulated at the post-transcriptional level. MicroRNAs are small non-coding RNAs that mediate post-transcriptional silencing of genes involved virtually in all biological processes. However, little is known about the roles that microRNAs play in SULT2A1 expression. In this study, we used MiRanda and TargetScan to predict microRNA candidates that potentially bind to SULT2A1 mRNA. The minimum free energy (MFE) of microRNA-mRNA interaction was then calculated via RNAhybrid to evaluate the strength of the binding. We identified potential binding sites of hsa-miR-495-3p and hsa-miR-486-5p in 5'UTR and 3'UTR of SULT2A1 mRNA. The MFE of the microRNA-mRNA hybridization was all less than −20 kcal/mol, indicating a high likelihood of microRNA-mRNA interaction in the cells. Furthermore, we extracted RNA-seq and miRNA-seq data from The Cancer Genome Atlas (TCGA) and conducted Pearson correlation analyses of the levels of SULT2A1 mRNA and miRNA candidates. We found that hsa-mir-495 and hsa-mir-486 are inversely correlated with SULT2A1 at the expression level in 48 non-tumor human liver samples. Our integrative analyses provide a foundation for investigating the repressive regulation of SULT2A1 by miRNAs.

AR-BIC – 14

Transcriptional Responses Reveal Similarities between Preclinical Rat Liver Testing Systems

Zhichao Liu[1]*, **Brian Delavan[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

Toxicogenomics (TGx) is an important tool to gain an enhanced understanding of toxicity at the molecular level. Previously, we developed a pair ranking (PRank) method to assess in vitro to in vivo extrapolation (IVIVE) using toxicogenomic datasets from the Open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATEs) database. With this method, we investigated three important questions that were not addressed in our previous study: (1) is a 1-day in vivo short-term assay able to replace the 28-day standard and expensive toxicological assay? (2) are some biological processes more conservative across different preclinical testing systems than others? and (3) do these preclinical testing systems have the similar resolution in differentiating drugs by their therapeutic uses? For question 1, a high similarity was noted (PRank score = 0.90), indicating the potential utility of shorter term in vivo studies to predict outcome in longer term and more expensive in vivo model systems. There was a moderate similarity between rat primary hepatocytes and in vivo repeat-dose studies (PRank score = 0.71) but a low similarity (PRank score = 0.56) between rat primary hepatocytes and in vivo single dose studies. To address question 2, we limited the analysis to gene sets relevant to specific toxicogenomic pathways and we found that pathways such as lipid metabolism were consistently over-represented in all three assay systems. For question 3, all three preclinical assay systems could distinguish compounds from different therapeutic categories. This suggests that any noted differences in assay systems was biological process-dependent and furthermore that all three systems have utility in assessing drug responses within a certain drug class. In conclusion, this comparison of three commonly used rat TGx systems provides useful information in utility and application of TGx assays.

AR-BIC – 15

Host Factors and Drug Properties Associated With the Development of Chronic DILI

Kristin McEuen-Ashby[1,2], Ayako Suzuki[3], Huixiao Hong[2], Mary Yang[1], Elizabeth Pierce[1], Hong Fang[2], Minjun Chen[2]

1 Department of Information Science, University of Arkansas at Little Rock and University of Arkansas for Medical Sciences, Little Rock, Arkansas

2 Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, Arkansas

3 Division of Gastroenterology and Hepatology, Duke University, Durham, North Carolina

Drug-induced liver injury (DILI) is a clinically significant adverse reaction. Many drugs have the potential to cause DILI, and although most DILI cases resolve after discontinuation of the culprit medication, approximately 10% liver injury will persist for months and develop into a chronic disease. Little is known about the mechanisms driving the development of persistent DILI. In this study, we used 335 well-defined DILI cases from an international consortium to explore the host factors and drug properties associated with persistent DILI. We defined cases as acute or persistent based on injury type and alanine aminotransferase levels; cholestatic cases were categorized as persistent if alanine aminotransferase remained elevated for at least six months and hepatocellular or mixed cases were persistent if alanine aminotransferase remained elevated for three months or more. Cases that normalized within six months if cholestatic and within 3 months if hepatocellular or mixed were categorized as acute. A total of 225 drug properties were retrieved from the Liver Toxicity Knowledge Base and literature sources, and 84 host factors were collected from the clinical data. Of the 335 DILI cases, 275 were classified as acute and 73 as persistent. We found that maximum bilirubin, maximum ALT, age, and sex were significantly associated with the development of persistent DILI. Of the 69 culprit drugs involved, 28 caused one or more persistent cases and the others only caused acute cases. Drug properties that were significantly associated with persistent drugs included mitochondrial liability, the number of hydroxyl groups, and estrogen receptor activity. The factors associated with persistent DILI need to be further validated.

AR-BIC – 16

Mining Pharmacogenomic Information from Drug Labeling Using FDALabel Database for Advancing Precision Medicine

Darshan Mehta[1], Ryley Uber[1], Zhichao Liu[1], Shraddha Thakkar[1], Minjun Chen[1], Joshua Xu[1], Baitang Ning[1], Shashi Amur[2], Padmaja Mummaneni[2], Steve Harris[1], Guangxu Zhou[1], Leihong Wu[1], Taylor Ingle[1], Junshuang Yang[1], Weida Tong[1], Hong Fang[1]

1National Center for Toxicological Research, US FDA, 3900 NCTR Road, Jefferson, AR 72079
2Center for Drug Evaluation and Research, US FDA, Silver Spring, MD 20993

It is known that drug response can have significant interpatient variability. The relationship between drug response and genetic makeup of an individual/population, studied in the field of pharmacogenomics, is rapidly accelerating advancements towards precision medicine. The US FDA has included pharmacogenomic information in the labeling of approved drug products to improve drug safety and efficacy. In this research, we use FDALabel database for mining pharmacogenomic information from drug labeling and propose a novel classification scheme for distinguishing drug-biomarker pairs related to safety from those related to efficacy. FDALabel is a web-based application that allows users to perform customizable searches of about 95,000 labeling documents that include human prescription and over-the-counter (OTC) drugs. Using a set of 62 biomarkers obtained from a public FDA resource website, we queried FDALabel database to identify drugs with biomarker information. Using this approach, we identified 225 unique drugs with pharmacogenomic information in their drug labeling and a total of 289 drug-biomarker pairs. These drug-biomarker pairs were then classified into 4 categories as either causing adverse reactions, having dose-related information, targeted for a genetic indication, or simply informative. An analysis of the relationship between drugs and biomarkers revealed that the most frequently observed biomarkers are CYP2D6, G6PD, and CYP2C19 and the most frequently observed therapeutic areas are oncology, psychiatry, and infectious diseases. The labeling sections with the most occurrences of biomarkers were found to be Clinical Pharmacology, Clinical Studies, and Indications and Usage. The results presented in this research show the utility of FDALabel database in mining pharmacogenomic information from drug labeling. As new biomarkers are discovered and more information is added to drug labeling, FDALabel can help researchers to guide the advancement of precision medicine.

AR-BIC – 17

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 – 18

How Important is the Chief Data Officer for the Company?

Yu Nie, John Talburt

College of Engineering and Information Technology, University of Arkansas at Little Rock, Little Rock, AR 72204, U.S.A

Abstract: Dramatic growth of big data industry creates a new position, chief data officer (CDO). The CDO is responsible for data governance and is able to influence the management process of a firm. However, what are the CDO development stages, what roles and responsibilities of the CDO and what effects this emerging position brings to firm are not clear. We divide the CDO development history to 5 stages: Generation stage, Growth stage, Rapid development stage, Prosperous stage and Recession stage. We find the most important roles and responsibilities of the CDO. The roles are: Data Governor, Commercial Value Digger, Decision Maker, Coordinator and Big Data Concepts and Skills Promoter. The most important responsibilities are maximizing data value and dig more business value from data. In order to investigate the impact of the CDO on the firm performance, we collected 6 years performance data from treatment firms and 3 years performance data from control firms. The data indicates the improvement of treatment firms after the CDO appointment and shows the performance differences between treatment firms and control firms. The results indicate that the profit ratios of treatment firms would be significantly improved after the firms appointed CDOs, and that the profit ratios of firms with CDOs would be significantly higher than that of the control firms without CDOs. For the cost ratios, our findings provide some empirical evidences revealing some cost ratios of treatment firms after CDOs appointment are lower although one cost ratio, SGA/S, is higher for treatment firms after CDOs appointment.

AR-BIC – 19

Origin of the Genetic Code

Joseph Onyilagha, and Annette Craig-Hall

University of Arkansas at Pine Bluff.

How nature used only four nucleotides to build its proteins and forming the genetic code are intriguing. Three theories have been propounded about the origin of the genetic code, a) the stereochemical theory, which deals with codon assignments as determined by physico-chemical affinity between amino acids and the anticodons, b) the Coevolution theory, which hypothesizes that the code structure coevolved with amino acid biosynthesis pathways, and c) the Adaptive theory, also known as error minimization theory, which is based on natural selection to lessen the detrimental effects of point mutations and translation errors. Our research has two goals, a) to update the biosynthesis pathways of the 20 standard amino acids, and b) to ascertain whether metabolic pathways found in living organisms can serve as accurate guides to ancient evolutionary events. We utilized comparative computer platforms such as BLAST and Protein Databank among others in our efforts to uncover relationships of enzymes involved in the biosynthesis pathway. Our results are consistent with previous reports that show some distinctions between the “early” and “late” amino acids of the genetic code. For example, biosynthetic steps in many of the late amino acids are longer than those in the early ones, and longer steps suggest the involvement of many more enzymes. Again, synthesis of some late amino acids is not a “one-way traffic” from early members because some late amino acids can give rise to some early amino acids through well-defined pathways. This finding complicates some of the key assertions of the Co-Evolutionary theory. Example, there is no precursor-product relationship that connects Glycine (Gly), early amino acid and Threonine (Thr), late amino acid. In other words, Thr cannot be synthesized from Gly and vice-versa. Nevertheless, our results show new pathways that facilitate the biosynthesis of Thr from Gly and/or vice-versa. Several other such examples exist in our new updated pathways with respect to early versus late amino acids.

Project was funded by NSF/HBCU-UP

AR-BIC – 20

Whole Genome Resequencing to identify QTLs for ascites in chickens.

Alia Parveen, Shatovisha Dey, Christa Jackson, Nicholas Anthony, Douglas Rhoads

University of Arkansas, Fayetteville, AR

We are using whole genome resequencing to identify chromosomal regions associated with resistance or susceptibility to ascites, a form of pulmonary hypertension syndrome, meat-type chickens. Previous Genome Wide Association Studies (GWAS) based on Single Nucleotide Polymorphisms (SNPs) have identified regions on chromosomes 2, 9 and Z. Despite several GWAS and further genotyping, there are no reliable or potential markers for ascites phenotype. We have completed screening of Copy Number Variations (CNVs) and Single Nucleotide Polymorphisms in ascites resistant and susceptible birds from the relaxed, REL, line derived from a commercial elite broiler line. DNA samples from resistant and susceptible birds were purified, quantified and pooled in two pools of 10 DNAs from each phenotype for both genders. Eight pools (2 pools x 2 phenotypes x 2 genders) were generated. Each pool was submitted for bar-coded library generation, and 2x125 paired end reads on Illumina HiSeq 2500 and with 66X genome coverage. The sequence reads were mapped onto Galgal5 using Bowtie for initial CNV mapping cn.mops (R package). Further mapping to chromosomes were done using NGen and ArrayStar (DNAStar ver 13). So far we have identified two potential regions for CNVs and 31 regions for SNPs with potential association with ascites phenotype. CPQ gene on chromosome 2 and LRRTM4 gene on chromosome 22 have been validated for containing ascites QTLs. However, their exact role in ascites is yet to be discovered. Further, we will validate all regions using high-throughput SNP assays in a larger catalog of DNAs from additional commercial broilers.

AR-BIC – 21

Directed Genome Evolution to identify Staphylococcal pathogenicity genes for macrophage survival and killing.

Abdulkarim Shwani,[1] Sura Zaki,[1] Sohita Ojha,[1] Douglas Rhoads[1]

[1-] 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. This staphylococcal species has been isolated from the bone and blood of lame broilers at the University of Arkansas. Bacterial chondronecrosis with osteomyelitis (BCO) has also been successfully induced by administration of a chicken isolate of S. agnetis (isolate 908) in drinking water. BCO primarily affects the growth plate in the proximal femur and tibia, the fast-growing leg bones. When birds are reared on suspended wire flooring with administration of strain 908 in the drinking water (10E5 CFU/ml on days 20 and 21) lameness incidence is as high as 80% by 56 days of age. The annotated complete genome of strain 908 has been published. In previous work, to better understand the relationship between dairy cattle and broiler isolates, we obtained nine S. agnetis isolates from milk or mammary gland secretions (n = 7) and udder skin (n = 2) from the University of Missouri for sequencing (2 x 250 MiSeq) and de novo assembly (NGen ver 13, DNAStar). To trace phylogenetic relationships, we constructed phylogenetic trees based on multi locus sequence typing using sets of genes or whole genome distance comparisons. Included in this analysis were published genomes from NCBI for S. agnetis and the closely related species Staphylococcus hyicus as outgroup. The chicken isolate, strain 908, clustered with two of the cattle isolates, one of them is 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 isolate 908 survives phagocytosis by chicken macrophage, while isolate 1379 cannot. For Directed Genome Evolution (DGE), DNA from strain 908 was electroporated into strain 1379 and the transformed population passaged through chicken macrophage. This produced a culture that also kills chicken macrophage. We sequenced the genome of the survivors and identified 5 gene regions and the three S. agnetis 908 plasmids in the survivors. We are now repeating the DGE with less total DNA and with purified plasmid to further refine which genetic determinants are required for macrophage killing.

AR-BIC – 22

Template Based Data Integration and Prediction (TB-DIP)

Shorab S [1], **Mahanaz S[1],** Jin J[2], Emine B[2], Patrick AL[2], Orloff M[2]

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

[2]Department of Epidemiology, University of Arkansas for Medical Sciences, Little Rock, AR

BACKGROUND

In general, small-mid size research laboratories struggle with managing clinical and secondary datasets. In addition, faster dissemination, correlation and prediction of information from available datasets is always a bottle neck. To address these challenges, we have developed a Template Based Data Integration and Prediction (TB-DIP) tool which is a hybrid model combining advantages of both relational and NoSQL technologies that provides flexible data integration mechanism using template based relational methodology and data visualization, knowledge discovery using NoSQL: Graph database.

METHODS and FINDINGS

A pipeline for the analysis of the incoming data was developed to load data from heterogeneous source to Relational Database via subject area Templates. Data from Relational database is cleansed and fed into Graph database using data import tools provided by Neo4j for Knowledge discovery and Predictive analysis.

Use Cases: A set of 750 patients with any type of lung cancer diagnosis and their clinical facts, family and social history were loaded into system. It was discovered, "C34.90, Lung Adenocarcinoma" was the most common diagnosis among the patients and the majority of these patients family members has history of cancer. A second use case for comparing clinical facts with molecular genetic analysis was undertaken. Random 90 patients that had clinical diagnosis of Lung Adenocarcinoma were selected, We compared their demographics (Race) and mRNA expressions for genes OLFML3 and SP1 and we found African-American patients diagnosed with Adenocarcinoma had mRNA expression mostly under expressed compared to white population for both genes OLFML3 and SP1.

CONCLUSION:

TBDIP proved to be great tool in (1) Integrating data from heterogeneous sources (2) Knowledge discovery on clinical dataset (3) Comparing clinical facts with molecular genetic analysis. The findings that we have may be generalized to other clinical areas and settings.

KEYWORDS

Integration, Prediction, NoSQL: Graph database, genomic, lung cancer.

AR-BIC – 23

Analyses of genome quality scores across 120,000 genomes

Visanu Wanchai and David Ussery

Arkansas Center for Genomic Epidemiology & Medicine and The Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205

Shortly after the advent of second generation of genome sequencing technologies, there began an exponential growth in sequencing data in public genome repositories. This large amount of genomic data allows us to investigate and better understand the diversity in many groups of bacteria. However, the majority of available sequenced genome sequences are not complete, and include draft genomes of varying qualities. Here, we assessed quality scores for more than 120,000 bacterial genomes from National Center for Biotechnology Information (NCBI). Bacterial genomes with all level of completeness were collected and combined to make a non-redundant set of bacterial genomes. The genome quality score for each was calculated by four different measurements: assembly quality, number of rRNA and tRNA genes, and the occurrence of conserved functional domains. The analysis of quality scores across a set of more than hundred and twenty thousand genomes find that most (perhaps 80% or more) of the genomes are of acceptable quality for many uses. Although, some genome sequences are of very poor quality, in a few cases even for 'complete' genomes. These scores can be used as cut-offs to get a high-quality set of genomes for testing bioinformatics tools or improving the analysis.

AR-BIC – 24

Infer the in vivo Point of Departure with ToxCast in vitro Assay Data Using a Robust Learning Approach

Dong Wang

FDA National Center for Toxicological Research, Jefferson, AR UNITED STATES, Dong.wang@fda.hhs.gov

The development and application of high throughput in vitro assays is an important development for risk assessment in the 21st century. However, there are still significant challenges to incorporate in vitro assays into routine toxicity testing practices. We developed a robust learning approach to infer the in vivo point of departure (POD) with in vitro assay data from ToxCast and Tox21 projects. We utilized assay data from ToxCast and Tox21 projects to derive the in vitro PODs for several hundred chemicals. These were combined with in vivo PODs from ToxRefDB regarding the rat and mouse liver to build a high dimensional robust regression model. This approach separates the chemicals into a majority, well predicted set; and a minority, outlier set. Salient relationships can then be learned from the data. For both mouse and rat liver PODs, over 93% of chemicals have inferred values from in vitro PODs that are within ±1 of the in vivo PODs on the log scale (the target learning region, or TLR) and R2 of 0.80 (rats) and 0.78 (mice) for these chemicals. This is comparable with extrapolation between related species (mouse and rat), which has 93% chemicals within the TLR and the R2 being 0.78. Chemicals in the outlier set tend to also have more biologically variable characteristics. With the continued accumulation of high throughput data for a wide range of chemicals, predictive modeling can provide a valuable complement for adverse outcome pathway based approach in risk assessment.

AR-BIC – 25

Deep learning for food contamination detection

Leihong Wu, Zhichao Liu, Hongjian Ding, Tanmay Bera, Weida Tong, Joshua Xu*

FDA National Center of Toxicological Research, 3900 NCTR Rd, Jefferson AR 72079

Insect pests are often associated with food contamination and public health risks. Currently, pest fragments are recovered from food samples and then examined under microscope by analysts to assess the level of risk, which is time-consuming and highly experts relied. Development of an automated approach to assisting human experts is urgently needed in this area to improve food safety.

Deep learning technologies have been widely applied in image recognition and classification nowadays, however few deep learning approaches have been developed regarding to the food safety area. With that said, we set out to apply a deep convolutional network based on VGG-16 structure to recognize beetle remains in food products.

Since available collections of food contaminating beetle images are really limited, we directly borrowed the features trained by some other image deep learning networks. These features were passed through our customized network with four stacked fully connected layers, followed by a final “softmax” output layer for classification purpose. In this “proof-of-concept” study, we focused on the task to classify 15 common species of storage beetles. In details, for each species, 3000-7500 images were used as training dataset and another 1500 images per species are used for model evaluation.

As a result, the deep-learning network delivered promising cross-validation performance. Its overall accuracy in species identification was comparable with our earlier work of using other machine learning methods with 625 human-engineered features that were extracted from the images through complex image analysis algorithms. An obvious and important advantage of this deep learning approach is to bypass this step of image feature extraction through complex and designed algorithms. It can thus be readily extended to image analysis for other types of food contamination, e.g., rodent hairs. To the best of our knowledge, this is the first time that deep learning networks have been applied in food contamination detection and our approach demonstrated that deep learning is quite efficient and accurate in detecting food contamination such as storage beetles.

AR-BIC – 26

A machine learning approach for prediction of the benefit-risk profile with multiple time-to-event endpoints

Wei Vivian Zhuang, Ph.D.

Division of Bioinformatics and Biostatistics, National Center for Toxicological Research (NCTR), US Food and Drug Administration (FDA), AR 72079

Background: In a clinical trial or animal study that evaluates treatment effects on multiple time-to-event endpoints, subjects may experience benefits, harms, or both during the follow-up period. Strategies for analyzing the time to multiple endpoints include the standard analysis based on the time to a composite endpoint or the first event, individual component analysis on the time to each event, and multivariate analysis based on categorized outcomes. These strategies are suitable and powerful in many, but not all, studies with multiple time-to-event endpoints. Methods: We propose a personalized benefit-harm machine learning analytical tree approach (PBHAT) that explicitly outlines and predicts the benefit-risk profile of a treatment at the subject level based on multiple time-to-event endpoints. We sequentially classify the subjects in a clinically meaningful order from the least desirable to the most desirable, e.g. from death to survival with adequate clinical response and without adverse events. Conclusions: Our simulations show that PBHAT may help predict drug safety and the benefit-risk profile of a treatment under many circumstances. There are also challenges waiting to be solved to make the approach suitable and powerful under more circumstances.

Keywords: Subgroup Analysis, Survival Analysis, Benefit-Risk Analysis, Personalized Medicine, and Drug Safety Prediction