2019 Annual EAB Meeting

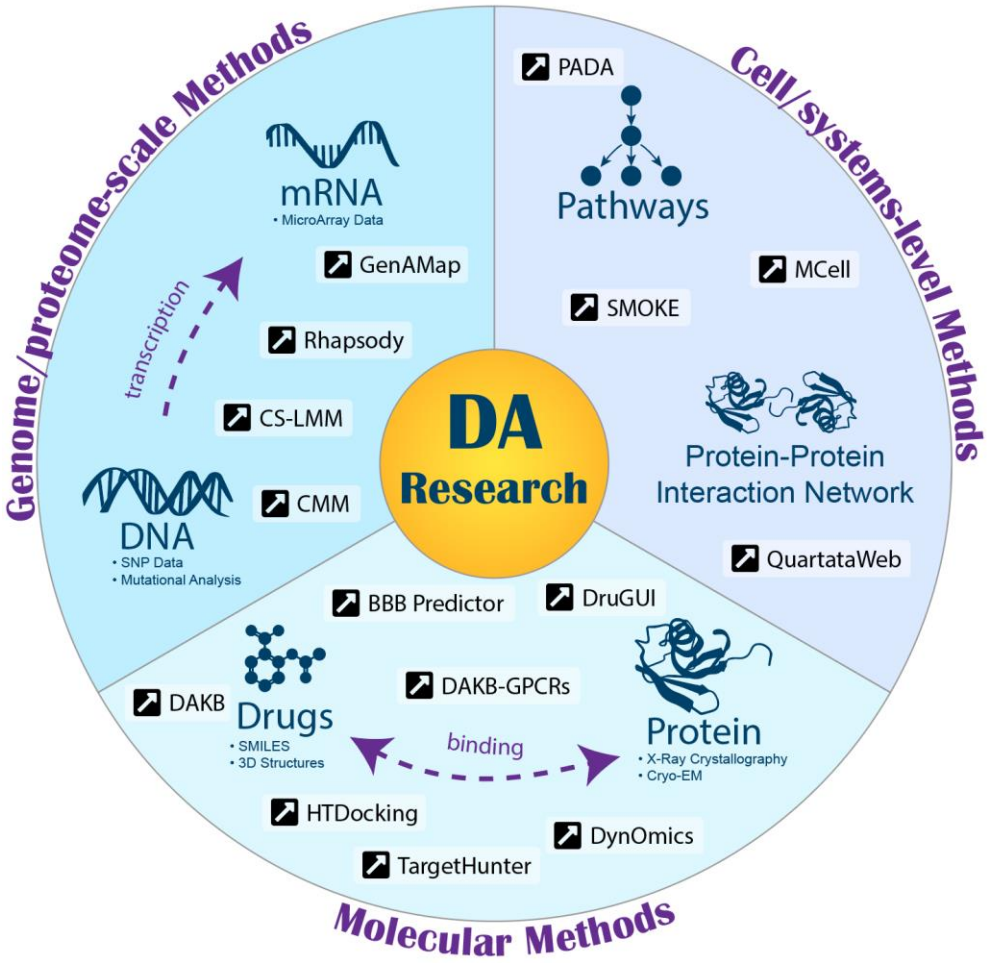
Skype-for-Business Video Conference

Thursday, October 17, 2019

NIDA-Funded National Center of Excellence for

Computational Drug Abuse Research (CDAR)

http://www.CDARcenter.org/



INDEX

NIDA National Center of Excellence for Computational Drug Abuse Research (CDAR)

(P30-DA035778)

Multiple PIs:

Xiang-Qun (Sean) Xie, PhD, EMBA
Director of National Center of Excellence for Computational Drug Abuse Research (CDAR)
Associate Dean of the School of Pharmacy
Professor of Pharmaceutical Sciences/Drug Discovery Institute
Director of Computational Chemogenomics Screening (CCGS) Center
University of Pittsburgh

Ivet Bahar, PhD
Distinguished Professor and John K. Vries Chair
Department of Computational & Systems Biology
School of Medicine, University of Pittsburgh

Eric Xing, PhD, PhD
Professor, Machine Learning Department
& Language Technology Institute & Computer Science Department
Associate Department Head of the Machine Learning Department
School of Computer Science, Carnegie Mellon University

OVERVIEW

The NIDA **Center of Excellence for Computational Drug Abuse Research** (www.CDARcenter.org, **CDAR**) is a joint initiative between the University of Pittsburgh (Pitt) and Carnegie Mellon University (CMU) under the leadership of Drs. Sean Xie and Ivet Bahar at Pitt and Eric Xing at CMU. The goal and mission of the CDAR Center during its first funding cycle has been to develop, implement and advance innovative technologies, and ensure the broad dissemination and usage of our data and tools to enhance and accelerate research in DA and related neurological disorders (ND), both in the local (Pittsburgh) area and nationwide. With these goals, the Center has become a leader in computational technology innovation and has successfully catalyzed synergistic collaborations between current and emerging researchers in the drug abuse (DA) research (DAR) area.

The high productivity of the Center during the past term is evidenced by **143 publications**, several of which were published in high profile journals. CDAR publications received >2,000 citations to date.

During the next funding cycle, the Center plans to further develop, integrate, and significantly augment the utility of our tools to enable DA-domain-specific chemical-to-protein-to-genomics-to-therapeutic intervention translation, being now equipped with a broader arsenal of tools and a larger number of collaborations with experts in DAR. The renewed CDAR Center will have three Research Support Cores (**Cores A-C**), and a new Core, **Core D**, for supporting the **Pilot/Feasibility Projects (F/PPs)**.

Core A, or the **Computational Chemogenomics** Core for DA, will help address polydrug addiction and systems pharmacological challenges by developing new chemogenomics resources and computing tools and by centralizing the data collected/generated by Cores A-C into a **Platform for Abused Drugs and Neurological Diseases Association (PANDA).** PANDA will be a DA- and ND-domains-specific chemogenomics repository and will serve as a national resource powered by GPU-accelerated cloud computing to enhance DA data dissemination, knowledge acquisition, and efficient use of computing technology among the broader DAR community.

Core B, or the **Computational and Systems Biology** Core, will develop methods and tools for mechanistic characterization of molecular-to-cellular events with special focus on neurotransmission events at the chemical synapses and systems-level responses to deficiencies in neurotransmission. In parallel, database-driven method and servers for designing effective intervention strategies will be developed which will take advantage of sequence, structure and pathway data. An innovative direction will be the development of a novel tool for assessing the pathogenicity of missense variants implicated in drug addiction.

Core C, or the **Computational Genomics** Core will synergize the work from Cores A and B by leveraging genome-scale approaches for genome-wide discovery of new DA targets and markers using machine learning and deep learning methods.

Cores A-C collaborate with **15 Funded Research Projects (FRPs)** centered around three themes: (i) substance of abuse-induced disorders and adaptations; (ii) DA-related disorders, cognitive syndromes and inflammatory diseases; and (iii) dopamine/neurotransmitter-signaling events/defects implicated in DA development and treatment.

The **Administrative Core** will continue to provide administrative support, using rules and policies established in the first term, in coordination with the **Scientific Steering Committee (SSC)** and consultation with an **External Advisory Board (EAB).**

Overall, the Center's overarching goal of translating advances in computational chemistry, biology, and genomics will accelerate the discovery of novel intervention methods for preventing, alleviating or treating drug addiction and associated NDs.

P30 EAB Annual Meeting Agenda* - October 17, 2019

(*All times listed below are EST. Locations of individual EAB members listed on page 5.)

3:00 – 3:10 pm **Video conference setup** *(CVS Room, Salk Pavillion, 335 Sutherland Dr.)* (*Light refreshments will be served.*)

3:10 – 3:20 pm **Welcome** by CDAR Center Dr. Sean Xie

Randall Smith, PhD, Professor and Senior Associate Dean, School of Pharmacy, University of Pittsburgh

Xiang-Qun (Sean) Xie, PhD, Director of CDAR Center & Professor and Associate Dean, Department of Pharmaceutical Sciences & Drug Discovery Institute, Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh

3:20 – 5:30 pm CDAR Core Research Reports by Core PIs (all team members invited)

3:20 – 3:55 pm Overview of CDAR Center: ***The First Funding Cycle & Beyond***
by **Dr. Xiang-Qun Xie**
AdminCore Report: ***Training, Dissemination and Administration***
by **Dr. Terry McGuire** (25 min + 10 minQ/A)

3:55 – 4:30 pm Core A Report: ***"Computational Chemogenomics & Systems Pharmacology for Drug Abuse and Neurodisorder Research"***

by **Xiang-Qun (Sean) Xie, MD, PhD, EMBA** (Core A), Professor and Associate Dean, Department of Pharmaceutical Sciences & Drug Discovery Institute, Director of CDAR Center (25 + 10 min Q/A)

4:30 – 5:05 pm Core B Report: ***"Molecular, Cellular and Systems Biology Methods and Tools for Computational Drug Abuse Research"***

by **Ivet Bahar, PhD** (Core B), Distinguished Professor & JK Vries Chair, Department of Computational & Systems Biology, Associate Director of University of Pittsburgh Drug Discovery Institute (25 + 10 min Q/A)

5:05 – 5:40 pm Core C Report: ***"Understanding Drug Abuse Using Deep Machine Learning Approaches"***

by **Eric Xing, PhD** (Core C), Professor, Machine Learning Department & Language Technology Institute & Computer Science Department, Associate Head of the Machine Learning Department, Carnegie Mellon University (25 + 10 min Q/A)

5:40 – 7:00 pm **Break (10 min) and Skype Meeting of EAB Members Only (70 min) (EAB Chairperson: Dr. Nurulain Zaveri)**

7:00 – 8:00 pm **Meeting of EAB members with the CDAR Leadership** (60 min)

EAB/SAB Members

Remote Skype Attendance Locations of EAB Members

Eric M. Billings, PhD – located in Bethesda, MD

Barry Gold, PhD – located in Crozet, VA

Tarek Leil, PhD – located in Princeton, NJ

Ying Mu, PhD, DABT – located in Laurel, MD

Chris Waller, PhD – located in Brookline, MA

Zheng-Xiong Xi, MD, PhD – located in Baltimore, MD

Nurulain T. Zaveri, PhD – located in Mountain View, CA

Eric M. Billings, PhD

Former Director & Staff Scientist, NIH NHLBI
Bioinformatics and Systems Biology Core Facility
Clinical Research Center, Bethesda, MD 20892
Email: ericbillings1@gmail.com

Dr. Eric Billings received his Physics B.S. degree from the University of California, Santa Cruz; Ph.D. in Biophysics from the School of Arts and Sciences, University of Connecticut in 1995; and post-doctoral training at the NIH as a National Research Council Fellow. Currently, he is a Staff Scientist in the Division of Intramural Research (DIR) at the National Institutes of Health. Dr. Billings is an expert in computational strategies to analyze and simulate biological processes. He has addressed biological questions ranging from the biophysics of enzyme catalysis to genetic drift in patient cohorts. His research focuses on integrating the analysis of the distinct data types found in biological systems.

Dr. Billings' post-doctoral training in CHARMM development centered on hybrid QM/MM molecular modeling methods. In order to provide the computational power necessary for this type of calculation, he and his post-doctoral mentor developed NIH's first commodity computer cluster, LoBoS (Lots of Boxes on Shelves). This system was the prototype for NIH's modern scientific cluster. During this time, he co-founded the NIH Molecular Modeling Interest Group.

He was the founding Director of the Genomics and Bioinformatics Core Facility which processed samples and analyzed data from RNA and proteomic microarrays, and partnered with Affymetrix' R&D to implement robotic sample processing for transcriptomic and genomic arrays. The Core developed novel methods to assess transcript profiling, gene regulation, regulatory motif discovery, mutational effects and perform genome wide association studies.

He was the founding Director of the Bioinformatics and Systems Biology Facility which supported intramural research requiring integration of 'Omics data with public data and prior knowledge. The facility developed methods for RNA analysis, meta-analysis of public data, pathway identification and a quantitative tool for the bench biologist to identify pathways of interest, simulate their behavior and compare model results to empirical data. During this time, he co-chaired the NIH Systems Biology Interest Group.

Prior to his scientific career, Dr. Billings worked in the computer industry. He joined a personal computer start-up company, Sirius Systems, and moved to Digital Equipment Corporation where he managed a team of software engineers supporting a Fortune 50 company. Dr. Billings has taught at American University and continues to mentor high school, college and post-baccalaureate students.

Barry Gold, PhD

Professor Emeritus, Department of Pharmaceutical Sciences

UNIVERSITY OF PITTSBURGH

Phone: 412 -480-5326

Email: goldbi@pitt.edu

Dr. Barry Gold was Professor and Chair of the Department of Pharmaceutical Sciences at the University of Pittsburgh. He also served as a co-Director of the University of Pittsburgh Drug Discovery Institute and was a member of the University of Pittsburgh Cancer Institute. Before moving to Pittsburgh, he was Professor and Associate Director for Basic Research at the University of Nebraska Medical Center's Eppley Cancer Center. He did his undergraduate, graduate and postgraduate studies in organic chemistry at Hunter College of the City University of New York, the University of Nebraska-Lincoln and the University of Toronto, respectively. His research interests are related to: (1) The design, synthesis and characterization of heterocyclic C-glycoside molecules that sequence-specifically bind to duplex DNA via triple helix formation. The goal of this work is to gene specifically regulation protein expression and probe transcriptional regulatory sequences; (2) The generation of anticancer drugs that efficiently form DNA lesions that are cytotoxic but not mutagenic in order to minimize secondary cancers; and (3) The development of small molecule inhibitors of specific DNA repair pathways that are involved in tumor resistance to anticancer drugs; and (4) The etiology on somatic mutations in tumor suppressor genes and oncogenes related to the development of cancer. He was active in the training of undergraduate, graduate and post-graduate students and taught courses in medicinal chemistry, biochemistry, biophysical biochemistry, chemical carcinogenesis, nucleic acids and cancer research.

Dr. Gold has published more than 140 peer-reviewed papers and been the lead inventor on three patents. He has been a member of the ACS since 1969. In 2008, he served as the Chair for the 31st National Medicinal Chemistry Symposium that was held in Pittsburgh. Previously, he has been on the executive committee of the ACS Division of Chemical Toxiciology. He has reviewed grants for a number of NIH study sections and for center and program project grants. He is on the editorial advisory boards of Burger's Medicinal Chemistry and Drug Discovery and Future Medicinal Chemistry, and serves as a scientific advisor to KeViRx LLC a small drug development company.Dr. Gold remains active, in efforts to increase the number of under-represented minorities in science. He has served as a mentor for high school teachers as part of the mentoring program of the Society for the Advancement of Native Americans and Chicanos in Science (SACNAS) and is on the Minority Affairs Committee of the Biophysical Society.

Tarek A. Leil, PhD

Group Director – Quantitative Clinical Pharmacology
Clinical Pharmacology & Pharmacometrics

BRISTOL-MYERS SQUIBB
Room E.2140C
Rt. 206 & Province Line Road
Princeton, NJ 08540-4000
Tel: (609) 252-6502
Email: tarek.leil@bms.com

Tarek Leil is currently the head of the Quantitative Clinical Pharmacology (QCP) group within Clinical Pharmacology and Pharmacometrics (CP&P) at Bristol-Myers Squibb (BMS). The QCP group at BMS uses model based approaches, including quantitative systems pharmacology (QSP), physiologically based PK (PBPK), and model-based meta-analysis (MBMA), to integrate clinical and non-clinical data using models that summarize knowledge and generate actionable predictions. These predictions can be used to help with decision making in drug discovery and clinical development, to optimize the design of patient clinical trials, and to facilitate communications with regulatory authorities. Prior to joining BMS in 2011, Dr. Leil worked at Pfizer in the Department of Clinical Pharmacology. Dr. Leil and the QCP group have numerous recent publications and presentations demonstrating the utility of Pharmacometrics and Systems Pharmacology in Pharmaceutical R&D.

Ying Mu, PhD, DABT

CEO

EagleImmune, Inc.
11361 Bishops Gate Ln
Laurel, MD 20723
Phone: 724-759-3066 (Cell)
Email: yingm6@gmail.com

Accomplished and board certified toxicologist at the FDA (former FDA officer) performed regulatory consulting and research, focusing on the evaluation of toxicological adverse events and biocompatibility of regulated products; conducting innovative research such as biomarker discovery and development through collaboration with internal and external scientists, specifically for allergic risks assessment comprised of two-arm approaches, *in vitro* and *in silico*, and that of clinical monitoring for applications in both premarket and post market; the effort is also intended to overcome the limitations of animal based tests toward the improvement of regulatory and clinical decision-making. As an effective Biocompatibility Standard Working Group member, Dr. Mu participated regulatory standard/guidance development with providing a broad array of knowledge in toxicology, biocompatibility, biomarkers, immunology, biomaterials, orthopedics, stem cell, tissue engineering, molecular biology, pre-market risk assessment and post-market safety evaluation, regulatory review process and the politics. Multilingual: Chinese, Japanese and English.
The founder of EagleImmune, started up in Maryland in June 2017.

Chris L. Waller, PhD

Vice President-Business Consulting and Chief Scientist
EPAM Systems, Inc.

M: 860-501-7113

Dr. Chris L. Waller is Vice President and Chief Scientist at EPAM Systems, Inc.

Dr. Waller was previously employed as Executive Director-Scientific Modeling Platforms by Merck and Co, Inc., Senior Director-Enterprise Architecture, Senior Director-Health Care Informatics, and Senior Director-Chemistry Informatics by Pfizer, Inc., and Director-Cheminformatics by Eli Lilly and Co., Inc.

He holds an Adjunct Full Professor position in the School of Pharmacy at the University of North Carolina-Chapel Hill where he lectures on big data, analytics, and drug design, is part-time faculty at Northeastern where he lectures on leadership in the College of Professional Studies, and serves as an advisor on data science related issues to the Hub at Davidson College, a data science incubator.

He is a founding member and serves on the board of the Pistoia Alliance, a life sciences industry pre-competitive collaboration consortium.

He received his undergraduate degree from Davidson College and a doctorate degree from the University of North Carolina – Chapel Hill.

Zheng-Xiong Xi, MD, PhD

Staff Scientist (2)
Intramural Research Program

NATIONAL INSTITUTE ON DRUG ABUSE, NIH
Phone: 443-740-2517 (Office), 443-668-3711 (Cell)
Email: ZXi@intra.nida.nih.gov

I am a senior Staff Scientist (level 2), Chief of Addiction Biology Unit, Molecular Targets and Medications Discovery Branch, National Institute on Drug Abuse (NIDA), Intramural Research Program (IRP). In the past 30 years, my research has been focused on brain mechanisms underlying drug reward and addiction and on mechanism-based medication development for treatment of addiction. The major expertise is regarding behavioral pharmacology, neurochemistry and optogenetics. Rat and mouse self-administration, neuroimaging, optogenetics, and *in vivo* microdialysis with HPLC are the core experimental techniques in the lab. The major findings include: 1) the classical inhibitory neurotransmitter receptor – GABAA receptor may also act as an excitatory neurotransmitter in the brain in some cases in 1990s; 2) identification of mGluR2 as a major glutamate autoreceptor modulating presynaptic glutamate release in 2000s; 3) discovery of brain cannabinoid CB2 receptor as a new potential target in medication development for treatment of addiction in 2010s; 4) identification of several addiction-related biomarkers (reduced D3 receptor or mGluR2 availability, polymorphisms at the T394 phosphorylation site of mu opioid receptor) in recent studies; and 5) discovery of several non-addictive phytocannabinoids such as cannabidiol, beta-caryophyllene and delt9-THCV may be promising medical cannabinoids for the treatment of substance use disorders. This research work has led to over 100 publications, including many of them published in such high-impact journals as Nature Nanotechnology, Nature Neuroscience, Cell Reports, PNAS, Neuropsychopharmacology, Journal of Neuroscience, etc. They are cited by over 5000. Some of them are reported by such news media and magazines as *Discovery*, *Time*, etc.

Nurulain T. Zaveri, PhD

Founder, President and Chief Scientific Officer
Principal Investigator

ASTRAEA THERAPEUTICS
320 Logue Avenue
Suites 120/131/142
Mountain View, CA. 94043
Phone: 650-254-0786 (Office)
Email: nurulain@astraeatherapeutics.com
Website: www.astraeatherapeutics.com

Dr. Zaveri is the Founder, President and Chief Scientific Officer of Astraea Therapeutics, a preclinical-stage company she founded in 2009, whose mission is medication development for under-served diseases of the central nervous system (such as substance abuse and addiction, chronic pain and Parkinson's disease). A PhD medicinal chemist by training, Dr. Zaveri is a seasoned pharmaceutical leader and recognized expert in the field of G-protein coupled receptor-targeted- and ion channel-targeted drug discovery for CNS medications. Pioneering discoveries from her laboratory have contributed to novel target validation for substance use disorders and pain treatment, and are being advanced into medication development. Dr. Zaveri has been a leader in the discovery and rational design of nociceptin opioid receptor ligands, under development for substance abuse treatment and chronic pain. Dr. Zaveri's discovery of the first truly high affinity and selective compounds targeted to the alpha3beta4 nicotinic ion channel receptors garnered tremendous interest in the nicotine addiction research arena and are being developed as smoking cessation medications. Before her entrepreneurial venture at Astraea Therapeutics, Dr. Zaveri was Principal Investigator and Director of the Drug Discovery Program at a nonprofit research institute for 16 years. Dr. Zaveri has been the Chair of the Drug Discovery and Development Interface (DDDI) Section of the American Association of Pharmaceutical Scientists (AAPS) and is an AAPS Fellow. Dr. Zaveri also serves on several NIH grant review committees. Dr. Zaveri is the lead inventor on over 15 patents and has authored over 66 research publications and 10 reviews/book chapters in fields of her research.

SHORT BIOGRAPHIES OF PIS

ABSTRACTS FOR P30 CORE A-C

&

CORE A-C HIGHLIGHTS

CORE A PI – SHORT BIOGRAPHY

Xiang-Qun (Sean) Xie, MD, PhD, EMBA

Associate Dean of the School of Pharmacy
Professor of Pharmaceutical Sciences/Drug Discovery Institute
Director of National Center of Excellence for Computational Drug Abuse Research and Computational Chemogenomics Screen Center
School of Pharmacy, University of Pittsburgh

Phone: 412-383-5276
Email: Xix15@pitt.edu

Dr. Xiang-Qun (Sean) Xie is Associate Dean for Research Innovation of the School of Pharmacy and Professor of Pharmaceutical Sciences/Drug Discovery Institute. He is a Director/PI of NIH National Center of Excellence for CDAR (www.CDARCenter.org). He is also a Founding Director of Computational Chemical Genomics Screening (CCGS) Center (www.CBLigand.org/CCGS) and a PI of an integrated Medicinal Chemistry Biology Laboratory of CompuGroup, BioGroup and ChemGroup. In addition, he serves as a Charter Member of the Science Board to the US FDA. He holds joint positions at Dept. of Computational Biology and Dept. of Structural Biology, and Pitt Cancer Institute MT/DD Program. He is an Editorial Advisory Board member for *AAPS Journal.* In the past, he served as a Chartered Member of NIH BPNS Study Section, and an *ad hoc* expert reviewer for UK MRC foundation; the Wellcome Trust Fund; the Netherland Organization for Scientific Research Council; the Austrian Science Fund (FWF); and the Chinese Natural Science Foundation.

Xie team is known for its pioneering research on development of diseases-specific chemogenomics knowledgebases, a platform of "Big Data to Knowledge" computational chemogenomics target identification and system pharmacology for translational research with over 150 publications and 10 patents. His work on Alzheimer's disease database was on ACS ***JCIM*** cover page story. The innovative work includes the GPU-accelerated cloud-computing machine-learning TargetHunter© program for drug target identification (***AAPS J*** special issue). His recent publications include computing technology development as illustrated: i) DAKB-GPCRs. An Integrated Computational Platform for Drug Abuse Related GPCRs (***J Chem Inf Model*** 2019**,** *59* (4), 1283-1289); ii) Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder (***Drug and Alcohol Dependence*** 2019 DOI:10.1016/j.drugalcdep.2019.107605); iii) Prediction of Orthosteric and Allosteric Regulations on Cannabinoid Receptors Using Supervised Machine Learning Classifiers (***Mol Pharm*** 2019**,** *16* (6), 2605-2615); and iv) Deep convolutional generative adversarial network (dcGAN) models for screening and design of small molecules targeting cannabinoid receptors. (***Mol. Pharm*.** 2019. DOI: 10.1021/acs.molpharmaceut.9b00500); as well as the applications as illustrated below: i) Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder (***Drug And Alcohol Depend*** 2019); ii) Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation (Drugs R & D 2019, 19 (3), 297-305); iii) Insight of captagon abuse by chemogenomics knowledgebase-guided systems pharmacology Target Mapping Analyses (***Sci Rep*** 2019, 9 (1), 2268); iv) Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose (ACS Chem Neurosci 2019, 10 (8), 3486-3499); v) the first discovery of INK4C chemical inhibitors for stem cell expansion (***Nature Comm*** 2015 and ***Scientific Report*** 2015); the first discovery of Sequestome-1 ZZ domain chemical inhibitors with therapeutics potential for tumors and neurological diseases (***Nature Leukemia*** 2016; ***Nature Comm***, 2017, PNAS 2018); and vi) over 50 cannabinoid publications of discover/report novel cannabinoid receptor CB2 functional ligands for cocaine attention, kidney fibrosis, multiple myeloma and osteoporosis (JMC 2013, ***Mol. Carcinog*** 2015, Kidney Ubt, 2918). He was a recipient of 2014 AAPS Award for Outstanding Research Achievements.

Dr. Xie received his MD in Pharmacy from the Second Military Medical University in Shanghai China, his PhD in Medicinal Chemistry from the School of Pharmacy University of Connecticut, and his EMBA degree from the School of Business Administration, University of Connecticut, USA.

CORE A: ABSTRACT

Computational Chemogenomics & Systems Pharmacology for Drug Abuse and Neurodisorder Research

*__Xiang-Qun (Sean) Xie,__ MD, PhD, EMBA, Junmei Wang, PhD, Lirong Wang, PhD, Zhiwei Feng, PhD, Ying Xue, PhD

National Center of Excellence for Computational Drug Abuse Research (CDAR);
Department of Pharmaceutical Sciences/Drug Discovery Institute,
Computational Chemical Genomics Screening (CCGS) Center, School of Pharmacy;
University of Pittsburgh, Pittsburgh, PA 15261, USA.

Core A, i.e., Computational Chemogenomics Core for Drug Abuse (**CC4DA**), has as its goal to address existing fundamental challenges in drug abuse (DA)/neurodisorders and medication research by systematically investigating interactions among chemical compounds and DA targets (proteins and signaling pathways). During the 1st funding year, we have advanced and expanded our developed computational algorithms and tools. We have made a magnificent progress to achieve the goals in the first funding cycle by publishing **58** peer-reviewed papers (**41** are directly related to DA and DA-associated NDs). All of these enhance the effectiveness of ongoing research and collaborations with the selected funded-research projects (**FRPs**) and also stimulate feasible pilot projects in the general realm of DA and neurological disorders. Under the leadership of PI Xie:

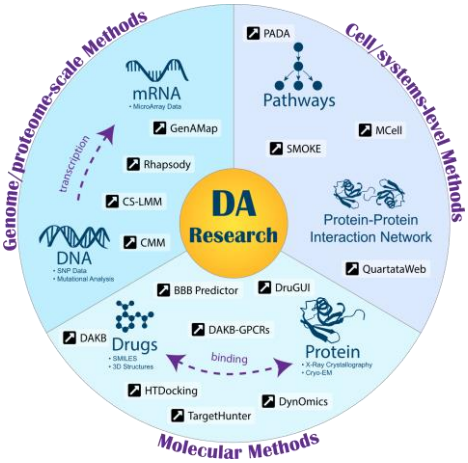
- We (or Core A) have continued to enrich the developed online chemogenomics knowledgebases in drug abuse (DA) and DA-associated neurological disorders. An alpha version of a new integrated computer platform - **Platform of Abused-Drugs and Neurological Diseases Association (PANDA)** has been constructed to integrate multiple (chemogenomics) databases and advanced computational technologies from all Cores to address the challenging problems in (poly)drug addiction and to collectively advance DAR by supporting the FRPs, guiding the P/FPs, and serving the broader DA/ND community.
- We have advanced our established new computational platform **DAKB-GPCRs[1]** and algorithms and tools (e.g., computational systems pharmacology-target mapping[2] and machine learning methods[3-5] for modeling abused drugs (e.g. opioids,[6] Captagon[7]) and DA target interactions (DTI), and to predict DA treatments against polyaddiction (e.g., fentanyl-laced cocaine overdose[2]).
- We have developed a set of ML-based models to predict DA clinical outcomes in combination with systems pharmacology method for assisting in DA prevention.[3-4] We are utilizing physiology-based pharmacokinetics (PBPK) modeling and pharmacometrics approaches to quantitatively study drug-drug interactions (DDIs) between opioids and benzodiazepines.[8]

Selected Publications:

1. Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X. Q., **DAKB-GPCRs: An Integrated Computational Platform for Drug Abuse Related GPCR**s. *J Chem Inf Model* **2019,** *59* (4), 1283-1289.
2. Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X. Q.; Feng, Z., **Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose.** *ACS Chem Neurosci* **2019,** *10* (8), 3486-3499.
3. Hu, Z.; Jing, Y.; Xue, Y.; Fan, P.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X. Q., **Analysis of substance use and its outcomes by machine learning: II. Derivation and prediction of the trajectory of substance use severity.** *Drug and Alcohol Dependence* **2019**, In press. DOI:10.1016/j.drugalcdep.2019.107605
4. Jing, Y.; Hu, Z.; Fan, P.; Xue, Y.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., **Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder.** *Drug And Alcohol Depend* **2019**, In Press. DOI:10.1016/j.drugalcdep.2019.107604
5. Bian, Y.; Jing, Y.; Wang, L.; Ma, S.; Jun, J. J.; Xie, X. Q., **Prediction of Orthosteric and Allosteric Regulations on Cannabinoid Receptors Using Supervised Machine Learning Classifiers.** *Mol Pharm* **2019,** *16* (6), 2605-2615.
6. Wu, X.; Xie, S.; Wang, L.; Fan, P.; Ge, S.; Xie, X. Q.; Wu, W., **A computational strategy for finding novel targets and therapeutic compounds for opioid dependence.** *PLoS One* **2018,** *13* (11), e0207027.
7. Wu, N.; Feng, Z.; He, X.; Kwon, W.; Wang, J.; Xie, X. Q., **Insight of Captagon Abuse by Chemogenomics Knowledgebase-guided Systems Pharmacology Target Mapping Analyses.** *Sci Rep* **2019,** *9* (1), 2268.
8. Ji, B.; Liu, S.; Xue, Y.; He, X.; Man, V. H.; Xie, X. Q.; Wang, J., **Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation.** *Drugs R & D* **2019,** *19* (3), 297-305.

CORE A: HIGHLIGHTS – 8/2014 - 8/2019

Construction of an integrated computer platform – PANDA. In conjunction with Core **B** and Core **C**, we have finished the construction of an alpha version of a computer Platform of Abused drugs and Neurological Diseases Association (**PANDA**). The platform integrates multiple (chemogenomics) databases and advanced computational technologies from **all Cores** to address the challenging problems in (poly)drug addiction and to collectively advance DAR by supporting the FRPs, guiding the P/FPs, and serving the broader DA/ND community.



The In-Development Platform of Abused- Drugs and Neurological Diseases Association (PANDA).

(1) Constructed a new platform DAKB-GPCRs and development of new algorithm "CSP-Target Mapping": **DAKB-GPCRs** is an Integrated computational platform for drug abuse related GPCRs.[1] And **"CSP-Target Mapping"**, is a new GPU-accelerated deep/machine learning-based algorithm tool for target prediction, classification, and mapping.[1-2] We applied our new platform and algorithm to study fentanyl-laced cocaine overdose[2] and Captagon abuse[3]. Our studies provided the potential abuse mechanism for these illicit drugs, and also provide the potential targets, off-targets, and therapeutics targets.

(2) Built predictive analytics models for DA prevention. In collaboration with R. Tarter, Director of CEDAR (CEDAR center was funded by NIDA), we constructed a set of ML/DL-based predictive models using the behavioral and psychological descriptors collected by CEDAR. The best model, which has achieved an area under the curve of receiver operating characteristic (ROC) of 0.71 using the data collected when the subjects are only at ages of 10 and 11. This model has a greater potential since substance abuse prevention measures will be taken when subjects are very young. So far two "**interesting and important**" [quote from editor] papers on drug abuse prevention have been accepted for publication in *Drug & Alcohol Dependence*,[4-5]

(3) Investigated drug-drug interactions (DDI) between opioids and benzodiazepines. We have conducted DDI studies using the population PK/PD and PBPK modeling in combination with molecular-level docking and physics-based binding free energy calculations. The PK parameters and covariates identified during population PK/PD analysis could be used to achieve precision medicine of DA treatment for individual patients. Our work studying DDI for opioid drugs has been published in *Drugs in R&D*[6] and another manuscript submitted.

References

1. Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X.-Q., DAKB-GPCRs: An Integrated Computational Platform for Drug Abuse Related GPCRs. ***Journal of chemical information modeling*** **2019,** *59* (4), 1283-1289. (PDF)
2. Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X.-Q.; Feng, Z., Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose. ***ACS chemical neuroscience*** **2019,** *10* (8), 3486-3499. (PDF)
3. Wu, N.; Feng, Z.; He, X.; Kwon, W.; Wang, J.; Xie, X.-Q., Insight of Captagon Abuse by Chemogenomics Knowledgebase-guided Systems Pharmacology Target Mapping Analyses. ***Scientific reports*** **2019,** *9* (1), 2268. (PDF)
4. Jing, Y.; Hu, Z.; Fan, P.; Xue, Y.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder. ***Drug And Alcohol Depend*** **2019**, https://doi.org/10.1016/j.drugalcdep.2019.107604 (PDF)
5. Hu, Z.; Jing, Y.; Xue, Y.; Fan, P.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X. Q., Analysis of substance use and its outcomes by machine learning: II. Derivation and prediction of the trajectory of substance use severity. ***Drug and Alcohol Dependence*** **2019**, https://doi.org/10.1016/j.drugalcdep.2019.107604 (PDF)
6. Ji, B.; Liu, S.; Xue, Y.; He, X.; Man, V. H.; Xie, X. Q.; Wang, J., Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation. ***Drugs R & D*** **2019,** *19*. (PDF)

CORE B PI – SHORT BIOGRAPHY

Ivet Bahar, PhD

Distinguished Professor & JK Vries Chair
Computational & Systems Biology Department,
School of Medicine, University of Pittsburgh

Phone: (412) 648-3332
Email: bahar@pitt.edu

Dr. Ivet Bahar is a Distinguished Professor, the John K. Vries Chair, and the Founding Chair in the Department of Computational & Systems Biology, at the University of Pittsburgh School of Medicine. She currently serves as Director/Principal Investigator of three multi-institutional NIH Centers (NIGMS-funded MMBioS National Center for Multiscale Modeling of Biological Systems, NHGRI-funded Big Data to Knowledge (BD2K) Center on Causal Modeling and Discovery (joint PI with Greg Cooper, Biomedical Informatics), and NIDA-funded CDAR (Computational Drug Abuse Research Center; Joint PI with Drs. Xie (Pitt Pharmacy) and Xing (CMU Machine Learning)); Associate Director, University of Pittsburgh Drug Discovery Institute; Founding Director (2005 – 2009), Executive Committee Member, and T32 co-PI of Carnegie Mellon University/Pitt PhD Program in Computational Biology.

Dr. Bahar is known as the developer of elastic network models for biomolecular systems dynamics, which found wide applications in molecular biophysics and computational and structural biology. Her current research areas include multiscale dynamics of biomolecular systems dynamics with applications to cell signaling, regulation and survival events (e.g. autophagy) in the central nervous system and immune system and neurodegenerative disorders; mechanisms of function of neurotransmitter transporters, evolution of proteins' sequence, structure, dynamics and function; protein-protein/ligand/lipid interactions, and supramolecular machinery and allostery; computer-aided drug discovery at both molecular and systems pharmacology levels. Dr. Bahar has co-authored in more than 250 papers in scientific journals, and her work has been cited more than 21,000 times to date (H-Index = 73).

Dr. Bahar is an elected member of the European Molecular Biology Organization (EMBO); Associate Editor of Proteins (Wiley); and editorial board member of scientific journals such as Structure (Cell Press), and Scientific Reports (Nature PBG). She regularly serves as a member of the NIH Biomedical Library and Informatics Review Committee (BLIRC) and is an invited speaker or keynote speaker at numerous international and national meetings, including European Molecular Biology Organization, Centre Européen de Calcul Atomique et Moléculaire, Gordon Research Conferences, and Annual meetings of the Biophysical Society and the American Chemical Society.

CORE B: ABSTRACT

Molecular, Cellular & Systems Biology Methods and Tools for CDAR

Ivet Bahar, Bing Liu, Mary H Cheng, Jiyoung Lee, James Krieger, Pemra Doruker & Fen Pei
Department of Computational & Systems Biology, School of Medicine, U of Pittsburgh

Recent years have seen a breakthrough in the elucidation of the structure and dynamics of membrane proteins. Sodium coupled neurotransmitter transporters are transmembrane proteins that are essential regulators of neurotransmission in the brain, and their malfunction is implicated in several neurological disorders. We have now made significant progress in understanding the complex machinery of these secondary transporters, the way they undergo cooperative structural changes between outward-facing and inward-facing states for transporting their substrate and sodium ions, while they also permit for chloride channeling[1]. We will present recent progress made in the elucidation of the mechanism of function of two major groups of transporters and their alteration by ligand binding and/or multimerization: Glutamate transporters, exemplified by the archaeal transporter Glt_{Ph} which served as a useful model for understanding the dynamics of excitatory amino acid transporters (EAATs); and dopamine transporters as an important member of transporters sharing the LeuT fold.[2-4] We used a combination of elastic network models and advanced molecular dynamics simulations to elucidate how the multidomain structure or multimerization properties are essential to altering not only their conformational dynamics, but also the coupled membrane remodeling in the synapse. We furthermore examined the allosteric dynamics of crucial signalling proteins including: lipoxygenase[5,6], PINK1[7], and KLF4[8], and assessed the absued drug associaed targets and pathways at systems level[9]. Results highlight the significance of adopting multi-scale approaches, in conjunction with experimental data for unravelling intermediates and inter- and intramolecular couplings that could not be otherwise inferred from static structures. Finally, we have developed two versatile web servers for analyzing shared signature dynamics tempered by local fluctuations[10] and automatic pharmacophore modeling[11].

Publications:

1. Cheng MH, Bahar I. (2019) Monoamine transporters: structure, intrinsic dynamics and allosteric regulation. *Nature Structural & Molecular Biology* 26, 545–556. (PDF)
2. Lee JY, et al. (2019) Druggability simulations and X-ray crystallography reveal a ligand-binding site in the GluA3 AMPA receptor N-terminal domain. *Structure* 27: 241-252. (PDF)
3. Ponzoni L, et al. (2018) Shared dynamics of LeuT superfamily members and allosteric differentiation by structural irregularities and multimerization. *Philos Trans R Soc Lond B Biol Sci* 373: 1749. (PDF)
4. Cheng MH, et al. (2019) Trimerization of dopamine transporter triggered by AIM-100 binding: molecular mechanisms and effect of mutations. *Neuropharmacology* [Epub ahead of print] (PDF)
5. Anthonymuthu T, et al. (2018) Empowerment of 15-lipoxygenase catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE. *J Am Chem Soc* 2018, 140 (51): 17835-17839. (PDF)
6. Mikulska-Ruminska K, et al. (2019) Characterization of differential dynamics, specificity, and allostery of lipoxygenase family members. *J Chem Inf Model*. [Epub ahead of print] (PDF)
7. Wang K, et al. (2018) PINK1 interacts with VCP/p97 and activates PKA to promote NSFL1C/p47 phosphorylation and dendritic arborization in neurons. *eNeuro* 5 (6) ENEURO.0466-18.2018 (PDF)
8. Zhou Z, et al. (2019). A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. (2019) *EBioMedicine* 19: 30312-30313. (PDF)
9. Pei F. et al. (2019) Quantitative systems pharmacological analysis of drugs of abuse reveals the pleiotropy of their targets and the effector role of mTORC1. *Front. Pharmacol.*, 10, 1-16. (PDF)
10. Zhang S, et al. (2019). Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity. *Mol Biol Evol*, 36(9), 2053-2068. (PDF)
11. Lee J, et al. (2019). Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations. *Protein Sci*. [Epub ahead of print]. (PDF)

CORE B – HIGHLIGHTS 8/ 2018 – 8/ 2019

During the 1st funding cycle, we published a total of 45 papers that acknowledged P30 DA035778 including those in *Nature Struct Mol Biol*[1], *Mol Biol Evol*[2], *eLife*[3], *Nucleic Acids Res*[4,5], *Cell*[6], *J. Amer Chem Soc*[7], and *Bioinformatics*[8,9]. Here we highlight a few in the period of August 2018 – August 2019.

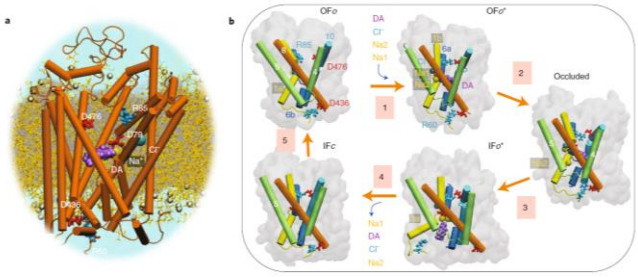


Fig 1. Structural dynamics of dopamine transporter. Details see ref (Cheng and Bahar, *Nat. Sturct. Mol Biol 2019*)

Elucidating the structural dynamics of monoamine transporters. Monoamine transporters (MATs) regulate neurotransmission via the reuptake of dopamine, serotonin and norepinephrine from extra-neuronal regions, thus maintaining neurotransmitter homeostasis. They are targets of antidepressants, substances of abuse and drugs for neuropsychiatric and neurodegenerative disorders. We elucidate structural dynamics of MATs (Fig 1) and their conformational landscape and transitions, and allosteric regulation mechanisms[1]. Oligomerization is a common feature of MATs. Yet, its effects on the function of MATs is not fully understood. In collaboration with the Amara and Sorkin labs (FRPs), we examined the possible mechanisms of human dopamine transporter (hDAT) dimerization[10] and trimerization[11]. Our study provides insights into mechanisms of addictive drug modulation.

Application of quantitative systems pharmacology (QSP) methods to drugs of abuse. We carried out a comprehensive analysis of cellular pathways implicated in a diverse set of 50 drugs of abuse using QSP methods[12]. Apart from synaptic neurotransmission pathways detected as upstream signaling modules that "sense" the early effects of drugs of abuse, pathways involved in neuroplasticity are distinguished as determinants of neuronal morphological changes. We found that many signaling pathways converged to mTORC1, which emerges as a universal effector of the persistent restructuring of neurons in response to continued substance abuse. Our analysis identified pathways enriched at different stages of drug addiction, as well as those implicated in cell signaling and regulation associated with DA.

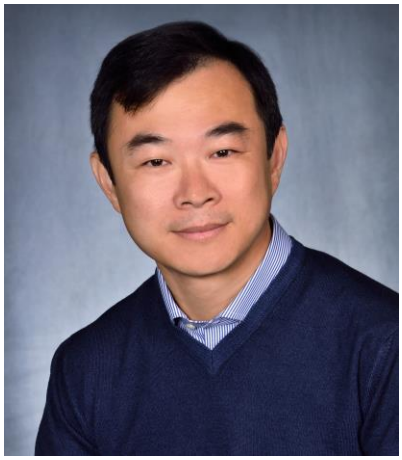
Regulation of ferroptosic cell death. In collaboration with the Wenzel lab (FRP), we discovered a key mechanism underlying the execution of ferroptosis[6,7]. We showed that phosphatidylethanolamine-binding protein 1 (PEBP1) serves as a regulator of ferroptosis upon binding 15-Lipoxygenases. PEBP1/15LO-driven ferroptosis occurs in asthma, kidney injury, and brain trauma modifying their activity and specificity to allow peroxidation of lipids. We demonstrated the importance of PEBP1-dependent regulation of ferroptotic death in cortical and hippocampal neurons in brain trauma. As master regulators of ferroptotic cell death with profound implications for human disease, PEBP1/15LO represents a new target for ND drug discovery.

Major achievements in method, tool and technology development. Our ***ProDy* API has reached a milestone of ~2 million downloads (as of August 2019; based on *Google Analytics).*** Core B has made major biotechnological advances during the past funding period: (1) We designed a new module Pharmmaker[13] for building pharmacophore model using outputs of druggability simulations (DruGUI). (2) QSP methods for predicting new drug-target associations, repurposable drugs and side effects have been implemented in BalestraWeb[8]. (3) We developed and implemented modules such as *Evol*[9] for bridging methods based on sequence evolution and structural dynamics to infer functional mechanisms. (4) We designed a new interface, DynOmics[5], for predicting and visualizing the environment-dependent dynamics of biomolecular systems. (5) We adapted the elastic network models (ENMs) originally introduced by our lab to new areas: predicting the effect of mutations (SAVs) on function[14], determining the conserved/specific dynamics of protein families/ subfamilies[2], and even chromosomal spatial dynamics[4]. These advances now open new opportunities to be realized by establishing and implementing new tools, RHAPSODY and *SignDy*.

References

(1) Cheng MH & Bahar I (2019) *Nat. Sturct. Mol Biol* **26**:545.(PDF)
(2) Zhang S, ... Bahar I (2019) *Mol Biol Evol* **36**:2053.(PDF)
(3) Cheng MH, …Bahar I (2017) *eLife* **6**:e25850.(PDF)
(4) Sauerwald N,…Bahar I (2017). *Nucleic Acids Res.* **45**:3663.(PDF)
(5) Li, H…. Bahar I. (2017). *Nucleic Acids Res.* **45**:W374.(PDF)
(6) Wenzel SE,…. Kagan VE (2017) *Cell*, **171**:628.(PDF)
(7) Anthonymuthu….Bayır H (2018), *JACS* **140:**17835.(PDF)
(8) Cobanoglu MC,… Bahar I (2015) *Bioinformatics* **31**:131.(PDF)
(9) Bakan A,... Bahar I (2014) *Bioinformatics* **30**:2681.(PDF)
(10) Cheng, MH,…Bahar, I. (2017) *J. Biol. Chem.***292:**12471.(PDF)
(11) Cheng, MH,…Bahar (2019). *Neuropharmacology*, in press.(PDF)
(12) Pei F,…Bahar I (2019). *Front. Pharmacol.*, **10:**1-16.(PDF)
(13) Lee JY, …Bahar I (2019) *Protein Sci, in press.*(PDF)
(14) Ponzoni L & Bahar I (2018) PNAS, **115:**4164.(PDF)

CORE c PI – SHORT BIOGRAPHY



Eric P. Xing, PhD

Professor
Machine Learning Department &
Language Technology Institute &
Computer Science Department,
Associate Department Head of the Machine Learning Department,
School of Computer Science
Carnegie Mellon University

Phone: (412) 268-2559
Email: epxing@cs.cmu.edu

Eric P. Xing is a Professor of Computer Science at Carnegie Mellon University. He is currently the Associate Department Head of the Machine Learning Department, founding director of the Center for Machine Learning and Health at Carnegie Mellon University, a Fellow of the Association of Advancement of Artificial Intelligence (AAAI Fellow), and an IEEE Fellow. He completed his undergraduate study at Tsinghua University, and holds a PhD in Molecular Biology and Biochemistry from the State University of New Jersey, and a PhD in Computer Science from the University of California, Berkeley. His main research interests are the development of machine learning and statistical methodology, and large-scale computational system and architectures, for solving problems involving automated learning, reasoning, and decision-making in high-dimensional, multimodal, and dynamic possible worlds in artificial, biological, and social systems. Prof. Xing currently holds or has held the following positions: associate editor of the Journal of the American Statistical Association (JASA), Annals of Applied Statistics (AOAS), IEEE Journal of Pattern Analysis and Machine Intelligence (PAMI) and the PLoS Journal of Computational Biology; action editor of the Machine Learning Journal (MLJ) and Journal of Machine Learning Research (JMLR); member of the United States Department of Defense Advanced Research Projects Agency (DARPA) Information Science and Technology (ISAT) advisory group. He is a recipient of the National Science Foundation (NSF) Career Award, the Alfred P. Sloan Research Fellowship in Computer Science, the United States Air Force Office of Scientific Research Young Investigator Award, the IBM Open Collaborative Research Faculty Award, as well as several best paper awards. Prof Xing is a board member of the International Machine Learning Society; he has served as the Program Chair (2014) and General Chair (2019) of the International Conference of Machine Learning (ICML).

CORE C – ABSTRACT

Understanding Drug Abuse using Deep Machine Learning Approaches

*__Eric Xing,__ PhD and Wei Wu, PhD

The goal of **Core C** (Computational Genomics Core for Drug Abuse, **CG4DA**) is to address fundamental methodological challenges of unraveling the genetic basis of DA by a systematic inference of the mapping between genetic variations and susceptibility to DA possibly induced by certain chemical compounds. Such a mapping provides a genome-wide atlas of potential targets and their risk under chemical compounds. During the first funding cycle, we have developed advanced machine learning (ML) approaches and software systems for drug abuse research. More important, we have made significant progress in establishing collaborations on the selected NIDA/NIH-funded research projects (FRPs) in the first term using the approaches and tools we developed. The following summarize our main achievements in this past year:

- To support the FRP PIs, Drs. Michael Vanyukov and Oscar Lopez, we developed machine learning models and methods, the Constrained Sparse Linear Mixed Model (CS-LMM) and the Coupled Mixed Model (CMM), to identify genetic variants jointly affecting substance use disorders and Alzheimer's disease. Using CS-LMM, we identified multiple SNP variants associated with alcoholism and/or AD.[1] Using CMM, together with Dr. Bahar and the Core B members, we identified five SNPs that are jointly associated with both SUDs and AD.[2]
- To support FRP PI Dr. Sally Wenzel's research, we developed a robust multiple kernel clustering method MML-MKKC which allowed us to cluster asthma patients using a wide variety of variables. We applied this approach to the 346 asthma patients, and identified four differential response patterns among the patients[3]. We also identified the top 12 baseline predictive variables, and validated the clusters using an independent SARP test set. These findings give insight into clinical, biologic and physiologic determinants of CS response patterns that could be mechanistically utilized to better link molecular to clinical responses.
- Following our previous work of detecting marginal epistasis signals, and motivated by the universal approximation power of deep learning, we developed a deep neural network method, namely Deep Mixed Model (DMM)[4], that can potentially model arbitrary interactions between SNPs in genetic association studies as an extension to the mixed models in correcting confounding factors. With simulations, we demonstrate the superior performance over the existing methods.

Selected Publications:

1. Wang H. Vanyukov MM, Xing EP, Wu W. Discovering Weaker Genetic Associations Guided by Known Associations, BMC Medical Genetics 2019. Accepted.
2. Wang H, et al. Coupled Mixed Model for joint genetic analysis of complex disorders from independently collected data sets: application to Alzheimer's disease and substance use disorder. Submitted to Bioinformatics. (https://www.biorxiv.org/content/10.1101/336727v2.article-metrics)
3. Wu W, et al. Multiview Cluster Analysis Identifies Variable Corticosteroid Response Phenotypes in Severe Asthma. American Journal of Respiratory and Critical Care Medicine 2019 Jun 1;199(11):1358-1367.
4. Wang H, Yue T, Yang Y, Wu W, Xing EP. Deep Mixed Model for Marginal Epistasis Detection and Population Stratification Correction in Genome-Wide Association Studies, BMC Bioinformatics 2019. Accepted.
5. Wu X, et al. A computational strategy for finding novel targets and therapeutic compounds for opioid dependence. PLoS One 2018;13(11):e0207027. eCollection 2018. PMID:30403753. PMCID: PMC6221321.
6. Bertholomey ML, et al. Phosphoproteomic analysis of the amygdala response to adolescent glucocorticoid exposure reveals G-protein coupled receptor kinase 2 as a target for reducing motivation for alcohol. Proteomes 2018. 6(4). pii: E41. PMID:30322021. PMCID: PMC6313880.
7. Wang H, Lengerich BJ, Aragam B, Xing EP. Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. Bioinformatics 2018. 35(7):1181-7.
8. Wang H, Aragam B, Xing EP. Variable selection in heterogeneous datasets: A truncated-rank sparse linear mixed model with applications to genome-wide association studies. Methods. 2018 Aug 1;145:2-9.
9. Wang H, Liu X, Xiao Y, Xu M, Xing EP. Multiplex confounding factor correction for genomic association mapping with squared sparse linear mixed model. Methods. 2018.
10. Marchetti-Bowick M, Yu Y, Wu W, Xing EP*. A penalized regression model for the joint estimation of eQTL associations and gene network structure. The Annals of Applied Statistics. 2019;13(1):248-70.
11. Wang H, et al. Graph-structured Sparse Mixed Models for Genetic Association with Confounding Factors Correction, BIBM 2019.

CORE C – HIGHLIGHTS

During the first funding cycle, we have published 41 papers acknowledging P30 DA035778. Many of these papers were published in high-impact clinical journals (e.g., Am J Respir Crit Care Med[1-3] and J Allergy Clin Immunol[4]), top machine learning and computational journals (e.g. Journal of Machine Learning Research,[5, 6] the Annals of Applied Statistics[7], Bioinformatics[8]), and the leading conferences for computational biology,[9, 10] and machine learning.[11-13]

Developing machine learning models and methods to identify genetic variants jointly affecting substance use disorders and Alzheimer's disease. Many genetic variants have individually small effects, but collectively large effects, on complex human diseases. These variants are difficult to discover using conventional statistical methods. In order to discover such variants associated with either alcoholism or Alzheimer's disease (AD) to support the PIs of the funded research project (FRP), Drs. Michael Vanyukov and Oscar Lopez, we developed a *Constrained Sparse Linear Mixed Model (CS-LMM)* (Aims 1 & 2). Using CS-LMM, we identified multiple potential weak but significant SNP variants associated with alcoholism and/or AD.[14] Motivated by these results, we recently developed another method, Coupled Mixed Model (CMM), that allow us to identify genetic variants jointly associated with two different types of diseases. Using CMM, we identified five SNPs that are jointly associated with both SUDs and AD.[15] One of the SNPs *rs224534* resides in *TRPV1*, which, as predicted independently by Dr. Bahar and the Core B members, is related to drug abuse. The paper describing these results is under review in Bioinformatics.

Developing machine learning methods for genome-wide association (GWA) mapping for marginal traits. We developed several methods for GWA mapping for marginal traits. Among them, a method named NETAM[9] showed promise on the 3-way association mapping among genome, transcriptome, and phenome. This method was developed in the first funding cycle to help the FRP PI Lopez better understand the genetic basis of AD. NETAM leverages transcriptome as a bridge between genome and phenome to boost the power of association mapping. We applied NETAM to an AD genetic data set which contained SNP data from 270 AD patients and 270 controls, and matching gene expression microarray data from prefrontal cortex, visual cortex and cerebellum of a subset of the subjects (GEO GSE44772). We found 477 associations from the AD data (Fig. 1), of which, 475 involve an SNP, an expression trait, and the phenotype (i.e., three-way associations). Only three of the 477 AD-associated SNPs were also identified by traditional analysis. To the best of our knowledge, AD-related 3-way associations have not been reported before. Notably, we found seven associations that involve VAMP1 encoding SNARE complex that controls neurotransmitter release via vesicle-mediated synaptic transmission; further, it is involved in nicotine pathway through SNARE complex. Interestingly, nicotine's involvement in AD has been extensively studied, and nicotinic receptors have been suggested as drug targets for AD. These findings appeared in the top computational biology conference ISMB.

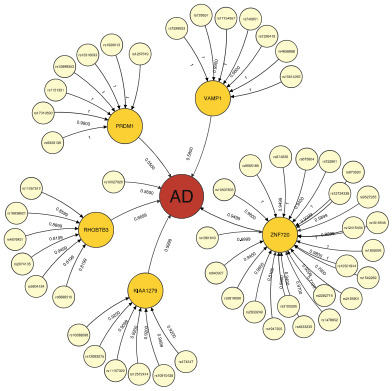


Fig. 1. *An example of 3-way association for AD identified by NETAM.*

Understanding complex response patterns of asthma patients to corticosteroids (CSs). CSs are the most effective asthma therapy, but responses are heterogeneous and systemic CSs lead to long-term side effects. In order to support FRP PI Dr. Sally Wenzel's research, we developed a robust multiple kernel clustering method MML-MKKC which allowed us to cluster asthma patients using a wide variety of variables. We applied this approach to the 346 participants in the Severe Asthma Research Program (SARP), and identified four differential response patterns among the patients[1]. We also identified the top 12 baseline predictive variables, and validated the clusters using an independent SARP test set. These findings give insight into clinical, biologic and physiologic determinants of CS response patterns that could be mechanistically utilized to better link molecular to clinical responses. Previously, we also helped Dr. Wenzel identify complex phenotypes among asthma patients.[2-4]

1. Wu W, Bang S, … Wenzel SE (2019) AJRCCM *199:* 1358. (PDF)
2. Modena BD, … Wu W …Wenzel SE (2014) AJRCCM 190: 1363. (PDF)
3. Modena BD, … Wu W …Wenzel SE (2017) AJRCCM 195: 1449. (PDF)
4. Wu W, … Wenzel SE (2014) JACI, 133: 1280. (PDF)
5. Al-Shedivat M, … Xing EP (2017) JMLR 18:1. (PDF)
6. Kolar M, … Xing EP (2014) JMLR 15:1713. (PDF)
7. Marchetti-Bowick M, … Xing EP (2019) *Ann. Appl. Stat.* 13:248. (PDF)
8. Marchetti-Bowick M, … Xing EP (2016) Bioinformatics 32:2903. (PDF)
9. Lee S, … Xing EP (2016) Bioinformatics (ISMB) 32:i164. (PDF)
10. Lee S, … Xing EP (2016) J Comput Biol (RECOMB) 23:372. (P
11. Xie P, … Xing EP (2017) ICML. (PDF)
12. Xie P, … Xing EP (2017) ICML. (PDF)
13. Xie P, … Xing EP (2018) ICML. (PDF)
14. Wang H, … Xing EP (2019) BMC Medical Genetics (Accepted
15. Wang H, … Xing EP (2019) Submitted to Bioinformatics. (PD

ADMINISTRATIVE CORE

- **AIMS**
- **ACCOMPLISHMENTS**
- **OPPORTUNITIES FOR TRAINING AND PROFESSIONAL DEVELOPMENT**
- **RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST**
- **PLAN FOR NEXT PERIOD**
- **PRODUCTS**
  - **PUBLICATIONS**
  - **WEBSITE(S) AND OTHER INTERNET SITE(S)**
  - **TECHNOLOGIES OR TECHNIQUES**
- **BIBLIOGRAPHY**

ADMINISTRATIVE CORE

For Computational Drug Abuse Research (CDAR) Center

Administrative Core Leaders: Sean Xie, Ivet Bahar and Eric Xing

EXECUTIVE SUMMARY
Xiang-Qun (Sean) Xie, MD, PhD, EMBA

PERSONNEL

Name	Organization	Role on Project
Xie, Xiang-Qun, PhD	University of Pittsburgh	PI/Director, Core A PI, and AdminCore PI
Bahar, Ivet, PhD	University of Pittsburgh	PI and Core B PI
Xing, Eric, PhD	Carnegie Mellon University	PI and Core C PI
Junmei Wang, PhD	University of Pittsburgh	Co-I & Coordinator for Core A
Zhiwei Feng, PhD	University of Pittsburgh	Co-I for Core A
Ying Xue, PhD	University of Pittsburgh	Co-I for Core A
Hongying Cheng, PhD	University of Pittsburgh	Co-I & Coordinator for Core B
Wei Wu, PhD	Carnegie Mellon University	Co-I & Coordinator for Core C, Co-I for Core D (on behalf of Core C)
Wang, Lirong, PhD	University of Pittsburgh	Co-I for Core D (on behalf of Core A)
Bing Liu, PhD	University of Pittsburgh	Co-I for Core D (on behalf of Core B)
Ayoob, Joseph, PhD	University of Pittsburgh	Enrichment Program Coordinator
Terence McGuire, PhD	University of Pittsburgh	CDAR Center Scientific Administrator

AIMS

Aim 1. To provide leadership for the *internal* operations of the CDAR Center to ensure efficient integration and coordination of CDAR member activities and to maintain a balance between productivity, innovation, and service.

Aim 2. To promote collaborative research and synergistic interactions between CDAR Core members and CDAR-affiliated members comprised of FRP and P/FP investigators and new potential collaborators.

Aim 3. To initiate and support information exchange and data sharing with the broader DAR community and further scientific research interactions with NIDA-pertinent research at-large.

ACCOMPLISHMENTS

Activities related to Aim 1. To provide leadership for the* internal *operations of the CDAR Center to ensure efficient integration and coordination of CDAR member activities and to maintain a balance between productivity, innovation, and service.

The center has been fully functional and running smoothly for the third year. The CDAR Center is directed by Xiang-Qun (Sean) Xie, PhD, EMBA, who is in charge of overseeing all activities, including scientific, service, and training activities. The scientific leadership is shared among three PIs, Drs. Xie, Bahar and Xing, who serve as the PIs of the three Cores. The constituency of the Steering Committee, External Advisory Board, and Coordinator positions are summarized below.

(1) **CDAR Scientific Steering Committee (SSC):** Senior Scientific Core members along with selected FRP PIs serve as the SSC board members to provide guidance and oversight to CDAR activities (see **Table 1**, below).

Table 1. Scientific Steering Committee (SSC)

SSC Member	Primary Affiliation	CDAR Role
Xiang-Qun (Sean) Xie, PhD	Dept of Pharmaceutical Sciences, SOP, Pitt	PI, Director, Core A PI and AdminCore PI
Ivet Bahar, PhD	Dept of Computational & Systems Biol, SOM, Pitt	PI, Core B PI
Eric Xing, PhD	Machine Learning Department, SCS, CMU	PI, Core C PI
Junmei Wang, PhD	Dept of Pharmaceutical Sciences, SOP, Pitt	Core A co-I and Core A Coordinator
Zhiwei Feng, PhD	Dept of Pharmaceutical Sciences, SOP, Pitt	Core A co-I
Ying Xue, PhD	Dept of Pharmacy and Therapeutics, SOP, Pitt	Core A co-I
Mary Cheng, PhD	Dept of Computational & Systems Biol, SOM, Pitt	Core B co-I and Core B Coordinator
Lirong Wang, PhD	Department of Pharmaceutical Sci, SOP, Pitt	Cores D co-I and Core D Coordinator
Wei Wu, PhD	Computational Biology Department, SCS, CMU	Cores C-D co-I and Cores C-D Coordinator
Bing Liu, PhD	Dept of Computational & Systems Biol, SOM, Pitt	Core B and D co-I and Core D Coordinator
Joseph Ayoob, PhD	Dept of Computational & Systems Biol, SOM, Pitt	Enrichment & Outreach Program Coordinator
Mary Toregrossa, PhD	Department of Psychiatry, SOM, Pitt	FRP11 PI (Theme 1)[(a)]
Sally Wenzel, MD	Environmental and Occupational Health, Pitt	FRP13 PI (Theme 2)
Andreas Pfenning, PhD	Computational Biology Department, SCS, CMU	FRP14 PI (Theme 2)
Alexander Sorkin, PhD	Department of Cell Biology, SOM, Pitt	FRP6 PI (Theme 3)

[(a)]*Themes: (1) cocaine & opioids; (2) DA-related inflammatory diseases and cognitive syndromes; (3) Neurotransmission.*

(2) **Core Coordinators:**

- **Junmei Wang, PhD** for Core A
- **Mary Cheng, PhD** for Core B
- **Wei Wu, PhD** for Cores C and D (on behalf of Core C)
- **LiRong Wang, PhD** for Core D (on behalf of Core A)
- **Bing Liu, PhD** for Core D (on behalf of Core B)

(3) **Enrichment and Outreach Programs Coordinator:**
Joseph Ayoob, PhD is in charge of assisting the Core PIs in establishing productive ties with existing centers and institutes in the Pittsburgh area as well as those from outside institutions.

(4) **External Advisory Board (EAB):**

The EAB has consisted of 9 members (7 retained after 1st funding cycle, see below). They are nationally renowned, with experience incheminformatics, computational biology or computational genomics, and research focus on DA, neurobiology and/or pharmacology.

- **Eric Billings, PhD**
  Director, Staff Scientist, Bioinformatics and Systems Biology Core Facility
  NIH, NHLBI
- **Barry Gold, PhD**
  Professor Emeritus, Department of Pharmaceutical Sciences
  University of Pittsburgh
- **Tarek A. Leil, PhD**
  Group Director – Quantitative Clinical Pharmacology
  Clinical Pharmacology & Pharmacometrics
  Bristol-Myers Squibb
- **Ying Mu, PhD, DABT**
  CEO, EagleImmune, Inc.
- **Christopher L. Waller, PhD**
  Vice President-Business Consulting and Chief Scientist
  EPAM Systems, Inc.
- **Zheng-Xiong Xi, PhD**
  Scientist, Intramural Research Program
  NIH, NIDA
- **Nurulain Zaveri, PhD**
  Founder, President and Chief Scientific Officer
  Astraea Therapeutics

Core technologies and programs have been expanded and/or updated.
In the fifth year, we continued to make significant advances in developing cutting-edge tools and platforms that can be readily utilized for DA research by both our collaborators and the external research community. The key technologies are illustrated below:

Core A:

1. **Pharmacometrics & Systems Pharmacology for drug abuse research (XQ Xie, JM Wang, LR Wang, ZW Feng, R Bertz)**
   a) We have constructed robust models for predicting substance use disorder and substance use severity (a newly established substance use/drug abuse outcome).
   b) We made progress to elucidate the drug-drug interaction mechanisms for oxycodone co-administrated with diazepam.
   c) We made progress to study the PK profile of heroine by using population PK modeling.
2. **Computer-Aided Drug Design for drug abuse research (XQ Xie, JM Wang, LR Wang, ZW Feng)**
   a) We have developed a novel approach, ELIE, to accurate calculate binding affinity.
   b) We have evaluated 17 protein force fields on studying the aggregation mechanisms for Amyloid-beta
   c) We have developed a novel computational protocol for designing inhibitors of Amyloid-beta aggregation. We have also studied the Amyloid-beta aggregation mechanism through massive molecular dynamics simulations
   d) We have studied a set of protein systems, including MscL, NK1R, TLR2, EGFR, etc., through collaborations.
3. **Novel Platforms or Tools for Virtual Animal and Drug Abuse (XQ Xie, ZW Feng et al)**
   a) We have finished the platform of virtual animal (MS is in prepar), which provide the experimental data/analyses of drugs for CVD. In addition, the platform can provide the prediction for the query compound(s)
   b) We have finished the platform of Drug Abuse Knowledgebase-GPCRs (MS in revision) that integrated with artificial intelligence algorithms

4. **Molecular mechanics force field development for drug abuse and neurological research. (JM Wang)**
    a) We have developed a set of atomic polarizability models for the new generation of AMBER polarizable force fields.
    b) We are about to finish the development of GAFF2.1, a milestone of GAFF force field development.
    c) We made substantial progress on redevelopment of a fast charge method (AM1-BCC) for rational drug design.

Core B:

1. We designed a new module **Pharmmaker** (http://prody.csb.pitt.edu/pharmmaker/) (*Protein Sci. in press*) for building pharmacophore model using outputs of druggability simulations (DruGUI). The pharmacophore models can be used for virtual screening of libraries of small molecules. A strong aspect of the method is that Pharmmaker uses multiple target conformations dependent on the binding poses of probes where they interact during druggability simulations, meaning that the binding score in virtual screening can be evaluated in a more realistic manner. Also, we can have multiple pharmacophore models with different target conformations and probe poses, which can be analyzed statistically.
2. We Implemented a new module ***SignDy*** (*Mol Biol Evol, in press*) to ***ProDy*** *API.* ***SignDy*** calculates the signature dynamics of families of proteins that share similar folds, but not necessarily similar sequences. Signature dynamics includes shared mode profiles, shared covariance between residue fluctuations, and their variations across family members. Additional information can be found in online tutorials;(http://prody.csb.pitt.edu/signdy/).
3. We developed an easy and efficient web server ***QuartataWeb*** (http://quartata.csb.pitt.edu) for mining known (experimentally verified) and predicted interactions for 5,494 drugs in DrugBank and 315,514 chemicals in STITCH, along with the confidence levels of the predicted chemical-target interactions (CTIs) using a machine learning based model.
4. We initiated the implementation of ***Rhapsody*** (http://rhapsody.csb.pitt.edu/) (*PNAS*, **115**: 4164-4169; 2018) for upgraded pathogenicity prediction of missense variants by taking structural dynamics into considerations.
5. significantly advanced the capabilities of ***ProDy***, which currently offers 10+ modules with user-friendly visualization tools, more than 40,000 code-line, and more than 4,000 pages of documentation including manuals and tutorial. *ProDy* reached an impressive milestone of 2 million downloads (http://prody.csb.pitt.edu/statistics/) as of September 2019.
6. Our database of GNM results, **iGNM DB (***Nucleic Acids Res* **44**: D415-422; 2016) now covers 95% of structures available in the PDB (a 5-fold increase compared to earlier version); and its improved techniques, libraries and markup language (Ajax, JQuery, HTM5, PHP and Highcharts) enhanced its security and interoperability.
7. Our webserver **DynOmics** (dynomics.pitt.edu) **(***Nucleic Acids Res,* **45:** W374-3*80;* 2017) is a portal developed to leverage rapidly growing structural proteomics data by efficiently and accurately evaluating the dynamics of structurally resolved systems, from individual molecules to large complexes and assemblies, in the context of their physiological environment.
8. Updated web server **Balestra (**http://balestra.csb.pitt.edu/) (*Bioinformatics* **31**:131-3**;** 2015) using the DrugBank version 5. Database architecture (PostgreSQL) has been used for the BalestraWeb server to improve the query performance. The searching engine has been improved to accept multiple proteins/drugs. We have integrated the protein information from the Uniprot database. We have improved the GUI/interface of the input and the output with the integrated information of proteins and drugs. The server has been extended to the Stitch database version 4 using the PostgreSQL database. We are developing the new version of BalestraWeb, which can efficiently identify chemicals, targets and pathways for drug abuse and will play an important role in discovering the underlying mechanisms and developing corresponding therapeutic strategies.
9. Implemented ***DruGUI*** (http://prody.csb.pitt.edu/drugui/) as a VMD plugin designed for setup and analysis of simulations containing small organic molecules (probes) for druggability assessment. *DruGUI* can incorporate a diverse set of molecules from CHARMM General Force Field (CGenFF) into simulations. *DruGUI* is used to i) identify *druggable* and *ligandable* sites; ii) setup a simulations that contain diverse probe molecules; iii) calculate probe molecule occupancy

grids; iv) analyze druggability of target protein; and v) perform druggability analysis of specific sites.

10. We developed **SMOKE** (https://liubing1020.github.io/smoke/) (*Autom Reas Syst Biol Med*, 63-92; 2019), a Statistical MOdel checKing tool for Estimating unknown parameters of dynamical models. It can utilize both quantitative data and qualitative knowledge for calibrating large models with hundreds of unknown parameters. It was originally developed for analyzing ordinary differential equation (ODE) models of biological networks, and currently being generalized to other modeling formalisms including stochastic models, rule-based models, and hybrid models.

Core C: **We have developed several new machine learning tools for DA research.**

1. Many genetic variants have individually small effects, but collectively large effects, on complex human diseases. These variants are difficult to discover using conventional statistical methods. In order to discover such variants associated with alcoholism and Alzheimer's disease, we developed a novel machine learning method called Constrained Sparse Linear Mixed Model (CS-LMM) (Aims 1-3). Using CS-LMM, we identified multiple potential weak but significant SNP variants associated with both alcoholism and/or Alzheimer's disease (AD).

2. Motivated by the results we observed from the CS-LMM project that some genes are associated with both alcoholism and AD, we extended the study to develop a new machine learning method, Coupled Mixed Model (CMM), that can identify genes that are jointly associated with two different types of diseases (Aims 1-3). In particular, we are interested in identifying the genes that are jointly associated with substance abuse disorder and AD by analyzing two independently collected data sets from raw sequence data.

3. Despite the proliferation of GWAS tools, detecting epistasis is still challenging. One main limitation of the existing tools is that they can only model linear association signals in the GWAS data. To overcome this challenge, we leverage the power of the deep neural networks and developed a tool, namely Deep Mixed Model (DMM), to model arbitrary interactions of the data.

4. Cluster analysis has been employed to detect subtypes of complex diseases which is a key task for precision medicine. However, clustering patients based on different sources/types of data (or called multiview data, e.g., clinical, gene expression, and proteomic data) can be challenging because different data has its own statistical property that is different from other data. Existing approaches that aim to address this problem can yield unfavorable results that largely depend on certain types of data when noise or redundant variables present in the multiview data. We developed a robust multiple kernel k-means clustering approach, called MML-MKKC, and showed that our method can robustly identify true clusters when noise or redundant variables are present in multiview data.

5. Better understanding how corticosteroids (CSs) use affects asthma patients is important for precision treatment of these patients. For this purpose, we analyzed a rigorously characterized adult asthma cohort from the Severe Asthma Research Program (SARP), and developed a multiview strategy which allows us to identify clusters of the asthma subjects with differential response patterns to CS using MML-MKKC. Using this strategy, we identified four clusters of patients showing differential response patterns among the asthma patients; our clusters were validated using an independent SARP test set.

Details are given in the reports from each individual core.

Activities related to Aim 2. *To promote collaborative research and synergistic interactions between CDAR Core members and CDAR-affiliated members comprised of FRP and P/FP investigators and new potential collaborators.*

A major goal of the NIH P30 PAR-18-225 is to provide research support to FRPs and P/FPs. Here, we present an overview of the 15 FRPs and 5 P/FPs selected for support in the second funding cycle (**Table 2, *top*** and **Table 3**, respectively). More details on Cores A-C activities to serve/accelerate/impact the ongoing FRPs and P/FPs are given in the reports from each individual Core. In addition, we selected four funded training programs (**FTPs**) led by Drs. Kass (CMU), Xu (Pitt), Rosano (Pitt) and Faeder/Bahar/Schwartz/Bar-Joseph (CMU/Pitt), to help train future DAR researchers (see **Table 2, *bottom***).

Table 2. NIDA-Funded Research Projects (FRP) that are supported by CDAR

#	PI (Institution)	Funding Source	Title	Status	Plan for year 6	Research Support Core
1	Venkat (Pitt)	5U54HD047905-15	BASIC/TRANSLATIONAL INVESTIGATIONS ON BUPRENORPHINE	On-going	Continue	A,C
2	Xi (NIDA)	1ZIADA12345	NOVEL CB2 FUNCTIONAL LIGANDS FOR COCAINE ATTENUATION	New FRP	Continue	A
3	Lopez (Pitt)	1RF1AG052525-01	SUBCLINICAL VASCULAR DISEASE AND ALZHEIMER'S DISEASE PATHOLOGY IN THE TRANSITION FROM MIDLIFE TO OLD AGE	On-going (1 paper published and 1 under revision)	Continue	A,C
4	Zhang (Pitt)	5R35GM128641-02	STRUCTURE, PHARMACOLOGY AND SIGNALING OF G PROTEIN-COUPLED RECEPTORS (GPCRS) IN INFLAMMATION	New FRP (*Formerly P/FP*) (1 paper under review)	Continue	**A**, B
5	Tang (Pitt)	1R01DA046939-01	ALPHA7 NICOTINIC RECEPTOR: STRUCTURES AND COUPLING WITH INTRACELLULAR PROTEINS	New FRP	Continue	**A**, B
6	Sorkin (Pitt)	2R01DA014204-15	REGULATION OF DOPAMINE TRANSPORTER (DAT) BY TRAFFICKING	On-going (4 papers published)	Continue	**B**
7	Newman (NIDA)	1ZIADA000389-22	NOVEL AND ATYPICAL DOPAMINE UPTAKE INHIBITORS	On-going (1 paper published)	Continue	**B**, A
8	Amara (NIMH)	1-ZIA MH002946	STRUCTURE, FUNCTION AND PHARMACOLOGY OF NEUROTRANSMITTER REUPTAKE SYSTEMS	On-going (*Formerly P/FP*) (1 paper published)	Continue	**B**
9	Chu (Pitt)	1R01NS101628-01	DENDRITE REGULATION BY THE MITOCHONDRIAL KINASE PINK1: IMPLICATIONS FOR	New FRP	Continue	**B**

			PARKINSON'S DISEASE/LBD			
10	Khoshbouei (U F)	5R01DA026947	METHAMPHETAMINE REGULATES THE DOPAMINE TRANSPORTER VIA AN INTRACELLULAR MECHANISM	New FRP	Continue	**B**
11	Torregrossa (Pitt)	5R01DA042029-02	MECHANISMS REGULATING COCAINE MEMORY STRENGTH	On-going (*Formerly P/FP*) (1 paper published)	Continue	**C**, B
12	Dong (Pitt)	2R37DA023206-12	COCAINE-INDUCED ADAPTATION IN NMDA RECEPTORS	On-going	Continue	**C**, A
13*	Wenzel (Pitt)	5P01AI106684-04	IMPACT OF INNATE AND ADAPTIVE IMMUNITY AT THE AIRWAY EPTHELIUM IN SEVERE ASTHMA	On-going (5 papers published)	Continue	**C**, **B**
14	Pfenning (CMU)	1DP1DA046585-01	INTERPRETING THE REGULATORY MECHANISMS UNDERLYING THE PREDISPOSITION TO SUBSTANCE USE DISORDERS	On-going (Formerly P/FP)	Continue	**C**, B
15	Alison Barth (CMU)	5R01NS088958-04	DYNAMIC CONNECTIVITY IN NEOCORTICAL NETWORKS	New FRP	Continue	**C**, B
NIH Funded Training Programs (FTPs) for CDAR Enrichment & Outreach Program Activities						
1	Kass (CMU)	5T90DA022762-12	INTERDISCIPLINARY TRAINING IN COMPUTATIONAL NEUROSCIENCE	On-going	Continue	A, B, C
2	Xu (Pitt)	5T32GM075770-12	RESEARCH TRAINING IN ANESTHESIOLOGY AND PAIN MEDICINE	New initiated	Continue	A, B, C
3	Rosano (Pitt)	1T32AG055381-01	POPULATION NEUROSCIENCE OF AGING AND ALZHEIMER'S DISEASE	New initiated	Continue	A, C
4	Faeder, Bahar (Pitt) Bar-Joseph (CMU)	2T32EB009403-11	INTEGRATED, INTERDISCIPLINARY, INTERUNIVERSITY PHD PROGRAM IN COMPUTATIONAL BIOLOGY	New initiated	Continue	A-C

* Although this project is NHLBI-funded, asthma research lies within the scope of NIDA-funded research.

Table 3. Pilot/Feasible Projects (P/FP) proposed for CDAR support in the 2nd funding cycle (a)				
P/FP	**PI**	**Institution**	**Title**	**Core**
1	Ying Xue	Pitt School of Pharmacy	SUD risk stratification by using ML algorithms and Bayesian network models	**A**, C
2	Blair Journigan	Marshall University	Novel ligands for TRPM8 menthol receptor for smoking cessation	**A**, B
3	Zachary Freyberg	Pitt Psychiatry & Cell biology	Cryo-EM approaches for drug abuse research	**B**, C
4	Min Xu	CMU Comp Biol	Deep learning approaches for analyzing Cryo-EM imaging data for drug abuse research	**C**, B
5	Scott Malec	Pitt-Biomedical Informatics	Literature-based discovery informing graphical causal modeling for repurposing drugs	**A**

(a) the Core that will lead the collaboration is written in boldface in the last column; as in Table 2.

Activities related to Aim 3. ***Further the CDAR's scientific interactions with NIDA-pertinent research at large, and to facilitate information exchange and data sharing with the broader community.***

This aim is achieved with the help of various enrichment and dissemination programs in the areas pertinent to DAR. Important activities as a national resource to benefit the broader community include:

- Outreach to the DAR community, contributions to existing NIDA-funded research and educational programs;
- Organization of workshops, courses, and online resources to enable efficient use of computational methods;
- Effective dissemination of software tools including begiining interactive tutorials, research progress and data, and maintenance CDAR web services and cloud computing server.

Since CDAR's inception in 2014, the AdminCore has played a critical role in establishing CDAR activities in local, regional, and national DAR communities. Under the leadership of the CDAR Director, and with the assistance of the Core Coordinators, the Enrichment and Outreach (**E/O**) Program leader (Dr. Joseph Ayoob), and the Scientific Administrator (Dr. Terence McGuire), the AdminCore has facilitated the dissemination and usage of state-of-the-art *in silico* tools and resources developed and maintained by Cores A-C. The AdminCore has also enabled the smooth operation and coordination of the complementary activities of the Cores as well as the productive collaborations with FRP and P/FP investigators and established DA researchers in academia and government agencies (NIH and FDA). **CDAR achievements** facilitated by the AdminCore **in the first funding cycle** include:

- Publications: **143**
- Conference Presentations: **77** *(see individual Cores for details)*
- Invited Talks by CDAR PIs: **67** *(see individual Cores for details)*
- Center Seminars: **65** (2018 – 2019 Seminars listed below; complete list at (http://www.cdarcenter.org/outreach/presentations/)
- Hands-on Training Workshops: **7 Courses** with **49 presentations** *(2018 – 2019)*
- CDAR-Hosted National Scientific Symposium: **1**

- **2** Patents and **1** Joint Disclosure (filed)
- CDAR-Supported Grant Submissions: **27**
    - **18** Awarded:
        - ❖ **5** to Junior Faculty
        - ❖ **6** NIH R01s
        - ❖ **2** Fellowships

Details are provided below and/or in the individual Core reports. See also CDAR website (http://www.cdarcenter.org).

Mentoring and Training Researchers in the Field

The training and course programs include courses and/or one-on-one work with a mentor, workshops, conferences, seminars, study groups, and individual study. The number of participants from the fist funding cycle is also listed, below (details are described in the report from each individual core).

a) PhD programs: PIs and co-investigators have been actively participating in the following **8** PhD programs:

a. PhD Program in Genomics, Proteomics and Drug Discovery (GPDD)
b. CMU/Pitt PhD Program in Computational Biology (CPCB)
c. PhD Program in Biomedical Informatics
d. Molecular Biophysics and Structural Biology (MBSB)
e. PhD Program in Medicinal Chemistry
f. PhD Program in Pharmacometrics and Systems Pharmacology (PSP)
g. PhD program in Machine Learning (CMU)
h. PhD program in Computer Science (CMU)

b) Teaching courses: The Core PIs and investigators have participated in the following **12** teaching courses:

a. Pharmacometrics & System Pharmacology
b. Drug Discovery, Design & Development Journal Club
c. Advanced Medicinal Chemistry
d. Foundations in Pharmaceutical Sciences
e. Pharmaceutical Analysis
f. Computational Systems Pharmacology (new course developed)
g. Graduate Machine Learning
h. Probabilistic Graphical Models
i. Computational Medicine
j. Genomics and Epigenetics of the Brain
k. Advanced Statistics
l. Advanced Pharmacokinetics
m. Computational Chemical Genomics for Drug Design

c) **Grad Students, Postdocs and Visiting Scholars Supervised: 60** *(first funding cycle)*
(See the report from each individual Core)
Core A – 32
Core B – 23
Core C – 5

d) Undergraduate Students Mentored: 15
(See the report from each individual Core)
Core A – 9
Core B – 5
Core C – 1

e) **Training Courses** – **7 Courses** with **49 presentations** *(2018 – 2019)*
In the last funding cycle, CDAR Center faculty have participated in numerous training courses/workshops (http://www.cdarcenter.org/outreach/fellows-training-workshop/). In the 2018 – 2019 period, the following courses and presentations were made by CDAR members:

August 6 – 29, 2019
10 presentations (4 classes) demonstrating the utilization of NONMEM for Systems Pharmacology - CDAR Workshop Series (Core A)

May 13-17, 2019
Hands-on Workshop on Computational Biophysics Pittsburgh Supercomputing Center. Pittsburgh, PA.
3 Presentations by Drs. Ivet Bahar, James Krieger, and Jiyoung Lee; Core B)

October 8, 2018 – April 11, 2019
32 presentations (12 classes) demonstrating the utilization of MATLAB and SimCYP for Systems Pharmacology - CDAR Workshop Series (Core A)

October 15-17, 2018
CECAM (Centre European pour le Calcul Atomique et Moléculaire) Workshop, "Multiscale simulations of allosteric regulatory mechanisms in cancer-associated proteins and signaling protein networks," Lugano, Switzerland. (presentation by Dr. Ivet Bahar).

October 6-8, 2018
Computational Biology Workshop, Arizona State University. Phoenix, Arizona (presentation by Dr. Ivet Bahar).

September 12-14, 2018
CECAM (Centre European pour le Calcul Atomique et Moléculaire) Workshop, "Normal modes of biological macromolecules: methods and applications," Paris, France (presentation by Dr. Ivet Bahar).

2018
National Research Mentoring Network (NRMN). Dr. Joseph Ayoob was one of four NRMN's Master Mentors, who was certified as a Train-the-Trainer Facilitator.

2016 Joint Core Hands-On Training Workshop *(part of P30 CDAR National Meeting)*
All three P30 Cores held a Joint Workshop for any interested attendees of our 2016 National Scientific Symposium meeting (August 9, 2016). Each Core presented 90 min talks highlighting the usefulness and proper application of their computational tools and gave practical demonstrations. An outline of the workshop is given below:

Core A Training (30 minute lecture and 1 hour demo & practice)
Training topics:
(1) ***Chemogenomics Database for Drug Abuse & Neuro-disorders:*** Genes, protein targets, pathways involved in a disease and small molecules that can directly interact with these key proteins with the potential to modulate the disease.
(2) ***TargetHunerMap:*** To predict the potential protein target(s) of a small molecule drug or a chemical compound through structure similarity or molecular docking.
(3) ***BBB prediction:*** To predict the blood-brain barrier **(BBB)** permeability of a small molecule based on its chemical features**.**

Core B Training (30 minute lecture and 1 hour demo & practice)
Training topics:
(1) ***ANM (Anisotropic network model):*** an elastic network based tool for analysis of dynamics of proteins and nucleic acids.

(2) *ProDy:* a free and open-source Python package for protein structural dynamics and sequence analysis.
(3) *Druggability simulations:* the use of DRUGUI, a VMD plugin, for performing MD simulations for druggability assessment, using probe molecules.

Core C Training (30 minute lecture and 1 hour demo & practice)
Training topic:
GenAMap: GenAMap software is a powerful platform for detection and easy visualization of structured association of genomic data and physical traits.

Outreach to students in existing training programs:

In the first funding cycle, students enrolled in other training programs were offered the opportunity to receive training with utilization of CDAR Center's tools as they may have application to their research projects (DA and DA-associated research). These training programs include:

- Interdisciplinary Training in Computational Neurosciences (R90DA023426, Dr. Kass, CMU)
- Research Training in Anesthesiology and Pain Medicine (5T32GM075770, Dr. Xu, Pitt)
- Program in the Neurobiology of Substance Use and Abuse (T32DA031111, Dr. Bradberry, Pitt)
- Population Neuroscience of Aging and Alzheimer's Disease (1T32AG055381, Dr. Rosano)
- Integrated, Interdisciplinary, Interuniversity PhD Program in Computational Biology (2T32EB009403, Drs. Faeder and Bahar (Pitt) and Dr. Bar Joseph (CMU))
- Joint Computational/Experimental Biomed Summer res Program for Undergraduate (R25DA032519, Dr. Madura, Duquesne)
- PhD Program in Genomics, Proteomics and Drug Discovery (GPDD, Pitt)
- CMU/Pitt PhD Program in Computational Biology (CPCB, Pitt/CMU)
- PhD Program in Biomedical Informatics (Pitt)
- Training and Experimentation in Computational Biology (TECBio) REU @ Pitt
- Molecular Biophysics and Structural Biology (MBSB, Pitt/CMU)

Publications

In the first funding cycle, the CDAR Center has developed to be a leader in technology innovation, successfully catalyzing synergistic collaborations between current and emerging researchers in the drug abuse (DA) research (DAR) area. The high productivity of the Center during the past term is evidenced by **143 publications**, several of which were published in high profile journals, such as *Nature Communications (IF = 11.9)*, *PNAS (IF = 9.6)*, *Cell (IF = 36.2)*, *Nature Structural & Molecular Biology (IF = 12.7)*, *Journal of Allergy and Clinical Immunology (IF = 12.5)*, *Current Opinion in Structural Biology (IF = 7.2)*, *eLife (IF =7.6)*, *Autophagy (IF = 11.1)*, *Nucleic Acids (IF = 11.1)*, *Alzheimer's & Dementia (IF = 14.4), American Journal of Respiratory and Critical Care Medicine (IF = 16.5), Journal of Medicinal Chemistry (IF = 6.3),* and *Molecular Biology and Evolution (IF = 14.8)*. The significance of the findings of CDAR Center's publications is reflected by the high number of times they have been cited by others in the field (**>2000 Google Scholar citations** as of September, 2019).

(The complete list of 143 P30 CDAR publications is provided in the PRODUCTS section at the end of the AdimnCore report, below.)

Invited Talks

In the first funding cycle, CDAR PIs presented **67** talks at national and international conferences. In the 2018 – 2019 period, CDAR Center PIs and senior members gave **24** invited talks.

(See individual Core reports for details.)

National Scientific Research Symposium hosted by CDAR Center – Aug. 8-9, 2016

In the first funding cycle, CDAR Center hosted a 2-day Scientific Research Innovation Research Conference which was held concurrently with our Second Annual P30 CDAR EAB Meeting and a joint Poster Session. The purpose of the symposium was to further the CDAR's scientific interactions with NIDA-pertinent research at large, and to facilitate information exchange and data sharing with the broader community. All Cores shared in the organization and presentation of research. Statistics gathered from the meeting are given below.

Summary of CDAR National Meeting Stats:

Meeting Stats

- **Total attendees:** 104
- **Major Talks:** 10 Total
  - ❖ 7 Invited Speakers:
    - 4 Academic (Northeastern U., CMU, Icahn Medical Institute, UPMC)
    - 2 Industry (Merck, Bristol-Myers Squibb)
    - 1 Government (NIH NIDA)
  - ❖ 3 Core PIs (Drs. Xiang-Qun Xie, Ivet Bahar, and Eric Xing)
- **Sponsors:** NIH NIDA, School of Pharmacy UPitt, Bristol-Myers Squibb, Dell, Walgreens, dotmatics
- **Total Students/Postdoc:** 20/15 (estimated)
- **No. of posters:** 21
- **3 Awardees:** *Cihan Kaya* (School of Medicine, UPitt), *Ziheng Hu* (School of Pharmacy, UPitt), and *Siwei Xie* (Computational Biology, CMU)
- **Government Agencies with Meeting Attendees**: FDA, NIH, NIDA, Pennsylvania Allegheny County Medical Examiner
- **Companies with Meeting Attendees**: Merck, Bristol-Myers Squibb, Astraea Therapeutics
- **Universities/Research Institutes with Meeting Attendees:** Northeastern University, Icahn Medical Institute, Albany College of Pharmacy and Health Sciences, UPMC, UPitt, CMU, Duquesne University, Fudan University (China), Tsinghua University (China), China Pharmaceutical University (China)

General Consensus of EAB Members: CDAR is in an excellent position to draw upon the strengths and research activities of three Principal Investigators (Drs. Xie, Bahar, and Xing), which possess complementary expertise and overlapping research objectives. The Center is well-positioned, with the combined objectives of the PIs, to enhance and implement their cutting-edge tools for identification and evaluation of targets; design and modeling of target ligands; and repurposing of existing drugs.

Training and Workshops Associated with National Symposium:

Our CDAR Center also held a Joint Workshop for any interested attendees of the meeting (August 9, 2016). Each Core gave 90 min talks highlighting the usefulness and proper application of their computational tools and gave practical demonstrations. An outline of the workshop is given below:

Core A Training (30 minute lecture and 1 hour demo & practice)
Training topics:

(1) *Chemogenomics Database for Drug Abuse & Neuro-disorders:* Genes, protein targets, pathways involved in a disease and small molecules that can directly interact with these key proteins with the potential to modulate the disease.

(2) *TargetHunerMap:* To predict the potential protein target(s) of a small molecule drug or a chemical compound through structure similarity or molecular docking.

(3) *BBB prediction:* To predict the blood-brain barrier **(BBB)** permeability of a small molecule based on its chemical features**.**

Core B Training (30 minute lecture and 1 hour demo & practice)
Training topics:

(1) *ANM (Anisotropic network model):* an elastic network based tool for analysis of dynamics of proteins and nucleic acids.

(2) *ProDy:* a free and open-source Python package for protein structural dynamics and sequence analysis.
(3) *Druggability simulations:* the use of DRUGUI, a VMD plugin, for performing MD simulations for druggability assessment, using probe molecules.

Core C Training (30 minute lecture and 1 hour demo & practice)
Training topic:

GenAMap: GenAMap software is a powerful platform for detection and easy visualization of structured association of genomic data and physical traits.

NOTE: Other training courses and hands-on workshops are discussed below.

OPPORTUNITIES FOR TRAINING AND PROFESSIONAL DEVELOPMENT PROVIDED

a) PhD programs: PIs and co-investigators have been actively participating in the following PhD programs:

a. PhD Program in Genomics, Proteomics and Drug Discovery (GPDD)
b. CMU/Pitt PhD Program in Computational Biology (CPCB)
c. PhD Program in Biomedical Informatics
d. Molecular Biophysics and Structural Biology (MBSB)
e. PhD Program in Medicinal Chemistry
f. PhD Program in Pharmacometrics and Systems Pharmacology (PSP)
g. PhD program in Machine Learning (CMU)
h. PhD program in Computer Science (CMU)

b) Teaching Courses: The Core PIs and investigators have participated in the following **9** teaching courses:

a. Drug Discovery, Design & Development Journal Club
b. Advanced Medicinal Chemistry
c. Foundations in Pharmaceutical Sciences
d. Pharmaceutical Analysis
e. Computational Systems Pharmacology (new course developed)
f. Graduate Machine Learning
g. Probabilistic Graphical Models
h. Computational Medicine
i. Genomics and Epigenetics of the Brain

c) Training Workshops: (described above)

d) Center Seminars. In the first funding cycle, CDAR Center has had a total of **65** seminars. Of these, in the 2018-2019 period, **24** scholars to gave presentations (see Table, below). The complete list of Center seminars can be found on the CDAR website (http://www.cdarcenter.org/outreach/presentations/).

Table 4. CDAR Center Seminars

Date	Lecturer	Topic
July 23, 2019	**Sarah Harris** Associate Professor Theoretical Physics Research Group University of Leeds	**Multiscale Simulations of Biological Macromolecules: from Atoms to the Continuum Limit**

April 16, 2019	**Minoli Perera, PharmD, PhD.** Associate professor Department of Pharmacology Northwestern University	**African Ancestry Pharmaco-omics**
April 16, 2019	**Lillian Chong, PhD** Associate Professor Dept. of Chemistry University of Pittsburgh	**The Art of Possibility: Weighted Ensembles of Trajectories**
April 2, 2019	**Dan Rosenbaum, PhD** Associate professor Departments of Biophysics and Biochemistry UT Southwestern Medical Center	**Structure and Mechanism of Human CNS GPCRs**
March 29, 2019	**Florencio Serrano Castillo** Graduate Student Researcher Chemical Engineering	**Multi-Scale QSP Models of Cystic Fibrosis Airway Pathophysiology**
March 26, 2019	**Patrick Marroum, PhD** Senior Director& Senior Research Fellow Clinical Pharmacology and Pharmacometrics AbbVie Inc.	**The Role of Modeling and Simulation in Setting Clinically Relevant Specifications**
March 19, 2019	**Timothy P. Ryan, PhD** Research Fellow Dyslipidemia and Outcomes Research Eli Lily and Company, Indianapolis, Indiana	**Incorporating Genomic, Clinical, and Real-world Evidence Into Drug Discovery-perspectives and Future Directions**
March 12, 2019	**Ali Saglam** University of Pittsburgh	**Weighted ensemble sampling of rule-based models**
March 5, 2019	**Jane M. Liebschutz, MD, MPH, FACVP** Chief, Division of General Internal Medicine Director, Center for Research on Health Care University of Pittsburgh School of Medicine	**Approach to Opioid Use Disorder in General Medical Settings**
February 26, 2019	**Yuanyuan Chen, PhD** Assistant Professor Dept. of Ophthamology University of Pittsburgh	**The Roles of Rhodopsin Homeostasis and Signaling**
February 12, 2019	**Weikang Wang, PhD** Postdoctoral Fellow Dept. of Computational and Systems Biology University of Pittsburgh School of Medicine	**Single cell morphology trajectory analysis on cell cycle and Epithelial-to-Mesenchymal transition**
November 20, 2018	**John Schuetz, PhD** Member, Vice Chair Department of Pharmaceutical Sciences St. Jude Children's Research Hospital	**ABC transporters as disease modifiers: Porphyria & cholestasis of pregnancy**
November 13, 2018	**Olivier Lichtarge, MD, PhD** Cullen Chair and Professor Molecular and Human Genetics Baylor College of Medicine	**Making Personal Sense of Disease: Machine Learning and Mutational Calculus**
November 6, 2018	**Yanqiao Zhang, MD** FAHA Professor of Integrative Medical Sciences Northeast Ohio Medical University	**Role of miR-34a in cardiometabolic diseases**
October 30, 2018	**Amy Newman, PhD *(New FRP)*** Deputy Scientific Director Medication Development Program Chief, Molecular Targets and Medications Discovery Branch NIDA-IRP/NIH	**Design of bitopic ligands with dopamine D2 receptor Bias**

October 30, 2018	**Luca Ponzoni, PhD** Postdoctoral Fellow Dept. of Computational and Systems Biology University of Pittsburgh School of Medicine	**Integrated Approach for Pathogenicity Prediction of Missense Variants**
October 30, 2018	**James Krieger, PhD** Postdoctoral Fellow Dept. of Computational and Systems Biology University of Pittsburgh School of Medicine	**Normal Mode Sampling Simulations and Electron Microscopy Image Analysis**
October 16, 2018	**David Boulton, PhD** Executive Director Quantitative Clinical Pharmacology AstraZeneca	**The role of clinical pharmacology in drug development: Dapagliflozin, a novel SGLT2 inhibitor, as an example**
October 2, 2018	**Kai Tan, PHD** Associate Professor Department of Pediatrics and Department of Biomedical and Health Informatics The Children's Hospital of Philadelphia University of Pennsylvania	**Optimal control nodes in disease-perturbed networks as targets for combination therapy**
May 22, 2018	**Ruiwen Zhang, MD, PhD, DABT, FAAAS** Professor of Pharmacology and Toxicology Robert L. Boblitt Endowed Professor in Drug Discovery Director of UH Center for Drug Discovery	**Targeting p53-MDM2 pathway: implication in cancer and neurogenerative diseases**
April 25, 2018	**Velvet (Blair) Journigan, PhD *(New P/FP)*** Assistant Professor of Medicinal Chemistry Department of Pharmaceutical Sciences Marshall University School of Pharmacy	**Structure-based design of novel small molecule ligands for the transient receptor potential melastatin 8 (TRPM8) ion channel: Insights and applications from recent structural biology advances**
March 21, 2018	**Mary Torregrossa, PhD *(Former P/FP, Now FRP)*** Associate Professor, Psychiatry, Translational Neuroscience Program, University of Pittsburgh	**Identification of Novel Regulators of Cocaine-Associated Memories**
January 8, 2018	**Inmaculada Hernandez, PharmD, PhD** Assistant Professor Department of Pharmaceutical Sciences University of Pittsburgh	**Real-world Use, Outcomes, and (Value-based?) Pricing of Pharmaceuticals: An Overview of My Research Program**
January 8, 2018	**Philip Empey, PharmD, PhD** Assistant Professor Department of Pharmaceutical Sciences University of Pittsburgh	**Pharmacogenomics to achieve precision medicine at Pitt/UPMC**

e) Joint Lab Meetings. In addition to the weekly meeting for all the cores, the Center also organizes monthly joint meetings between Core PIs and Core supported FRP Investigators. These are research work-in-progress sessions where students, postdocs, and faculty present informal research seminars, followed by thorough discussions. PIs of the P/F projects gave regular presentations for the evaluation of progress, with Center SSC members in attendance. These P30 CDAR meetings promoted knowledge sharing, research synergism, and building of collaborations. Below, we have listed the

talks that were presented during the 2018 – 2019 funding period. The complete list of joint monthly Meeting talks can be found at (http://www.cdarcenter.org/outreach/meetings/monthly/).

Invited Speakers for CDAR Center Monthly Meetings (2018 – 2019) – 13 Talks, 17 Presenters

July 10, 2019

Dr. Lirong Wang, Dr. Hongchun Cheng, Dr. Wei Wu and HaoHan Wang

Talk Title 1: "The integration of *Quartataweb* and *Target Hunter"* (Hongchun and Lirong)

Talk Title 2: "Genome-scale target and compound identification. The first prototype of an interface to be delivered by the end of June" (Lirong, HaoHan, and Wei)

Zhiwei and Ryan

Talk Title: "The PANDA Interface"

Junmei and Jiyoung

Talk Title: "The progress in *Druggability/Parameterization* Interfaces"

June 10, 2019

Leadership Meeting for P30 Resubmission

March 21, 2019

Dr. Xibing He and Beihong Ji

Talk Title: "Improve the Simulations of Biological Systems from Two Aspects: Force Field and Sampling"

Beihong Ji

Talk Title: "Phamacokinetics modeling and molecular modeling of Drug-Drug interaction between opioid and benzodiazepine."

Jan 24, 2019

Haohan Wang (Graduate Student, Eric Xing's lab, CMU)

Talk Title #1: "Coupled Mixed Model for Joint Genetic Analysis of Complex Disorders with Two Independently Collected Data Sets"

Talk Title #2: "Removing Confounding Factors Associated Weights in Deep Neural Networks Improves the Prediction Accuracy for Healthcare Applications"

December 3, 2018

Fen Pei (Core B Graduate student)

Talk Title: "Quantitative systems pharmacological analysis of drugs of abuse reveals the pleiotropy of targets and the effector role of mTORC1."

July 31, 2018

Leadership Meeting for P30 Resubmission

June 5, 2018

Zhiting Hu

Talk Title: "Text Generation: Algorithms and Toolkits"

May 22, 2018

*Ruiwen Zhang, MD, Ph.D., DABT, FAAAS***,** Professor of Pharmacology and Toxicology, Robert L. Boblitt Endowed Professor in Drug Discovery, Director of UH Center for Drug Discovery

Talk Title: "Targeting p53-MDM2 pathway: implication in cancer and neurogenerative diseases."

April 25, 2018

Velvet (Blair) Journigan, Assistant Professor of Medicinal Chemistry, Department of

Pharmaceutical Sciences, Marshall University School of Pharmacy
Talk Title: "Structure-based design of novel small molecule ligands for the transient receptor potential melastatin 8 (TRPM8) ion channel: Insights and applications from recent structural biology advances."

March 19, 2018

Mary Torregrossa (Associate Professor, Department of Psychiatry, Translational Neuroscience Program, University of Pittsburgh School of Medicine)
Talk Title: "Identification of Novel Regulators of Cocaine-Associated Memories."

f) Support of New Funding Efforts of FRP PIs: CDAR AdminCore also assisted in **27** CDAR-supported grant submissions (**18** awarded: **5** to junior faculties, **6** NIH R01s, and 2 fellowships) (see individual Core sections for details).

Table 5. Status of Research Grant submissions assisted by CDAR resources/activities (2014-19)

	Investigator (*Affiliation*)	Grant Title	Funding Source/Grant #
Funded Awards	Mary Torregrossa (*Pitt*)	Mechanisms regulating cocaine memory strength	NIH 5R01DA042029
	Andreas Pfenning (*CMU*)	Regulatory mechanisms underlying predisposition to substance use disorders	NIH 1DP1DA046585
	Oscar Lopez, D Sun (*Pitt*)	Chemogenomics syst pharmacology approach for TBI and AD research	DOD W81XWH-16-1-049
	Inmaculada Hernandez (*Pitt*)	Claims data mining to predict side effects of anti-dementia drugs	ALZ AARGD-17-500234
		Patient, system-level determinants of oral anticoagulation in atrial fibrillation	NIH K01HL142847
	Yong Wan (*Northwestern*)	Targeting the interplay between KLF4 and PRMT5 in carcinogenesis	NIH R01CA202963
	Satdarshan Monga (*Pitt*)	YAP & beta-catenin interactions in liver: implications in pathophysiology	NIH 5R01CA204586
		University of Pittsburgh liver research center	NIH P30KDK120531
	Seojin Bang (*CMU*)	Knowledge distilled multiview learning for identifying disease subtypes	CMU ML Fellowship
	David Perlmutter (*Wustl*)	New therapies for liver fibrosis and hyperproliferation in α1-AT deficiency	NIH P01DK096990
	Sweet Robert (*Pitt*)	Synaptic resilience to psychosis in Alzheimer Disease	NIH R01MH116046
	Joel Greenberger (*Pitt*)	Signature directed sequential delivery of small molecule radiation mitigators	NIH 5U19AI068021
	Rama Mallampalli (*Pitt*)	Immunosuppression in acute lung injury	NIH 2P01HL114453
	Alexander Sorkin (*Pitt*)	Regulation of dopamine transporter by trafficking	NIH 5R01DA014204
	Sally Wenzel (*Pitt*)	Type-2 or not type-2: this is the (therapeutic) question	NIH 5UG1HL139098
	Cheng Zhang (*Pitt*)	Structure, pharmacology and signaling of GPCRs in inflammation	NIH R35GM128641
	Ziv Bar-Joseph (*CMU*)	Comprehensive, infrastructure, mapping & tools for HuBMAP HIVE	NIH 1OT2OD026682
	Henry Dong (*Pitt*)	Myeloid FoxO1 in lipid metabolism	NIH 1R01DK120310
Pending	Zachary Freyberg (*Pitt*)	A multidisciplinary approach to decipher dopamine D2R signaling	NIH R35
	Valerian Kagan (*Pitt*)	Redox phospholipoxysome regulation of epithelial ferroptosis in asthma	NIH R01
	Hulya Bayir & Clark (*Pitt*)	Mitochondria-targeted therapies for cerebral ischemia in developing brain	NIH R01
	Sally Wenzel (*Pitt*)	Protein-oxidized phospholipid interactions in epithelial cell fate & asthma	NIH R01
Unfunded	Filippo Pullara (*Pitt*)	Comp sys pathology platform for analysis of hyperplexed imaging data	PITT CTSI Fellowship
	Stephanie Aldrich (*Pitt*)	Predicting resting- and active- structures of Cav2.1 Ca^{++} channel VSD III	PITT Mellon Fellowship
	X Lu, J. Wang (*Pitt*)	Integrative platform: causal inference, chemogenomics for target discovery	NIH U01
	Kevin Xiao (*Pitt*)	Novel biomarkers for atrial fibrillation (AF) severity	AHA Center Grant
	A. Van Demark (*Pitt*)	PEBP1 interactions regulating polyunsaturated phospholipid signaling in asthma	NIH R01

RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST

Dissemination of findings and sharing of data and copyright protected information are available through the CDAR Network over the Internet. Established resources, including hardware/software, are shared with others who have interests in applying computational technologies available to the

CDAR center to their research projects. Avenues whereby results are disseminated and information is exchanged include:

(1) Center Programs: all the cores have participated in interactive educational enrichment programs supported by the Center. They have been described, in part, above. Complete lists can be found at the CDAR Center's website (http://www.cdarcenter.org). They are briefly described below.

(a) CDAR Seminars. The Center has worked closely with Pitt and CMU departments (academic homes of the PIs) to select speakers relevant to the goals of the Center in general, and shared costs for invited speakers, selected from amongst leading scientists/laboratories in the field of DAR. In the third year, we have invited 17 scholars to give speeches in P30 meeting for knowledge sharing and communicating. (see details in section B4)

(b) Journal Clubs. All Cores also organize weekly literature review presentations in the areas of computational chemogenomics (Dr. Xie), computational biology (Dr. Bahar), and computational genomics (Dr. Xing) in addition to the departmental journal clubs. Trainees/students in each department are also required to make at least two presentations a year.

(c) Joint Lab Meetings. In addition to the weekly meeting for all the cores, the Center also organizes monthly joint meetings between Core PIs and Core supported FRP Investigators. These are research work-in-progress sessions where students, postdocs, and faculty present informal research seminars, followed by thorough discussions. (See details in section B4)

(d) Fellows Training Course/Workshop. (See details, above)

(e) CMU/Pitt Educational Opportunities. Our research has been disseminated by teaching in class. CDAR investigators have contributed to existing educational programs at Pitt and CMU, including the Computational Chemistry and computer-aided drug design lectures in the Advancing Medicinal Chemistry and Pharmaceutical Foundation courses. Dr. Ayoob works together with these programs' coordinators to enable appropriate contributions from CDAR PIs that will bring a higher profile to DAR and provide education, research, and training opportunities through mentored research opportunities and didactic components to students interested in this field of study.

(2) Scientific Progress & Dissemination of Results:

(a) Updated Core Technologies and Programs

(See Products section, below.)

(b) Conferences: During the first funding cycle, Center members have attended several conferences, both national and international) and presented:

(See individual Core reports for details.)

(c) Invited Talks: In the first funding cycle, CDAR PIs presented **67** talks at national and international conferences. In the 2018 – 2019 period, CDAR Center PIs and senior members gave **24** invited talks.

Core A – 10
Core B – 12
Core C – 2

(See individual Core reports for details.)

(d) Posters: In the 2018 – 2019 period, CDAR Center PIs and senior members gave **27** invited talks.

Core A – 14
Core B – 9
Core C – 4

(See individual Core reports for details.)

(e) National Scientific Research Symposium hosted by CDAR Center – Aug. 8-9, 2016

In the first funding cycle, CDAR Center hosted a 2-day Scientific Research Innovation Research Conference which was held concurrently with our Second Annual P30 CDAR EAB Meeting and a joint Poster Session. The purpose of the symposium was to further the CDAR's scientific interactions with NIDA-pertinent research at large, and to facilitate information exchange and data sharing with the broader community. All Cores shared in the organization and presentation of research.

PRODUCTS

Publications

As of September, 2019, the Center has a total of **143 publications**, 53% of which[1-75] are focused directly on either DA[1-46], DA-related health issues (i.e., asthma, cardiovascular problems)[47-59], or on associated neurological disorders (**ND**s)[60-75], all of which fall within the scope of the Program Announcement (PAR-18-225). Of the remaining publications, 52 describe development and implementation of methods and software, as we proposed in the 1st cycle proposal[76-127], and 16 describe the findings from collaborative studies with researchers using our resources[128-143]. It is worth noting that within the past year (since Oct 1st, 2018), we have **43** new publications, **59%** of which are focused on either DA or DA-related health issues. All 143 CDAR Center publications are listed below and are clustered in the 5 categories mentioned above.

I. Drug-abuse (DA) (46)

1. Alqarni, M.; Myint, K. Z.; Tong, Q.; Yang, P.; Bartlow, P.; Wang, L.; Feng, R.; Xie, X.-Q., **Examining the Critical Roles of Human CB2 Receptor Residues Valine 3.32 (113) and Leucine 5.41 (192) in Ligand Recognition and Downstream Signaling Activities.** *Biochem. Biophys. Res. Commun.* **2014,** *452* (3), 334-339.
2. Bertholomey, M. L.; Stone, K.; Lam, T. T.; Bang, S.; Wu, W.; Nairn, A. C.; Taylor, J. R.; Torregrossa, M. M., **Phosphoproteomic Analysis of the Amygdala Response to Adolescent Glucocorticoid Exposure Reveals G-Protein Coupled Receptor Kinase 2 as a Target for Reducing Motivation for Alcohol.** *Proteomes* **2018,** *6* (4), 41.
3. Bian, Y.; Feng, Z.; Yang, P.; Xie, X. Q., **Integrated In Silico Fragment-Based Drug Design: Case Study with Allosteric Modulators on Metabotropic Glutamate Receptor 5.** *AAPS J* **2017,** *19* (4), 1235-1248.
4. Bian, Y.; Jing, Y.; Wang, L.; Ma, S.; Jun, J. J.; Xie, X. Q., **Prediction of Orthosteric and Allosteric Regulations on Cannabinoid Receptors Using Supervised Machine Learning Classifiers.** *Mol Pharm* **2019,** *16* (6), 2605-2615.
5. Bian, Y. M.; He, X. B.; Jing, Y. K.; Wang, L. R.; Wang, J. M.; Xie, X. Q., **Computational systems pharmacology analysis of cannabidiol: a combination of chemogenomics-knowledgebase network analysis and integrated in silico modeling and simulation.** *Acta Pharmacol Sin* **2019,** *40* (3), 374-386.
6. Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X. Q., DAKB-GPCRs: **An Integrated Computational Platform for Drug Abuse Related GPCRs.** *J Chem Inf Model* **2019,** *59* (4), 1283-1289.

7. Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X. Q.; Feng, Z., **Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose.** *ACS Chem Neurosci* **2019,** *10* (8), 3486-3499.

8. Cheng, M. H.; Bahar, I., **Molecular Mechanism of Dopamine Transport by Human Dopamine Transporter.** *Structure* **2015,** *23* (11), 2171-81.

9. Cheng, M. H.; Bahar, I., **Monoamine transporters: structure, intrinsic dynamics and allosteric regulation.** *Nat Struct Mol Biol* **2019,** *26* (7), 545-556.

10. Cheng, M. H.; Block, E.; Hu, F.; Cobanoglu, M. C.; Sorkin, A.; Bahar, I., **Insights into the Modulation of Dopamine Transporter Function by Amphetamine, Orphenadrine, and Cocaine Binding.** *Front Neurol* **2015,** *6*, 134.

11. Cheng, M. H.; Garcia-Olivares, J.; Wasserman, S.; DiPietro, J.; Bahar, I., **Allosteric modulation of human dopamine transporter activity under conditions promoting its dimerization.** *J Biol Chem* **2017,** *292* (30), 12471-12482.

12. Cheng, M. H.; Kaya, C.; Bahar, I., **Quantitative Assessment of the Energetics of Dopamine Translocation by Human Dopamine Transporter.** *J Phys Chem B* **2018,** *122* (21), 5336-5346.

13. Cheng, M. H.; Ponzoni, L.; Sorkina, T.; Lee, J. Y.; Zhang, S.; Sorkin, A.; Bahar, I., **Trimerization of dopamine transporter triggered by AIM-100 binding: Molecular mechanism and effect of mutations.** *Neuropharmacology* **2019**, 107676.

14. Cheng, M. H.; Torres-Salazar, D.; Gonzalez-Suarez, A. D.; Amara, S. G.; Bahar, I., **Substrate transport and anion permeation proceed through distinct pathways in glutamate transporters.** *Elife* **2017,** *6*.

15. Dutta, A.; Krieger, J.; Lee, J. Y.; Garcia-Nafria, J.; Greger, I. H.; Bahar, I., **Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences.** *Structure* **2015,** *23* (9), 1692-1704.

16. Feng, R.; Tong, Q.; Xie, Z.; Cheng, H.; Wang, L.; Lentzsch, S.; Roodman, G. D.; Xie, X. Q., **Targeting cannabinoid receptor-2 pathway by phenylacetylamide suppresses the proliferation of human myeloma cells through mitotic dysregulation and cytoskeleton disruption.** *Mol Carcinog* **2015,** *54* (12), 1796-806.

17. Feng, Z.; Alqarni, M. H.; Yang, P.; Tong, Q.; Chowdhury, A.; Wang, L.; Xie, X. Q., **Modeling, molecular dynamics simulation, and mutation validation for structure of cannabinoid receptor 2 based on known crystal structures of GPCRs.** *J Chem Inf Model* **2014,** *54* (9), 2483-99.

18. Feng, Z.; Hu, G.; Ma, S.; Xie, X. Q., **Computational Advances for the Development of Allosteric Modulators and Bitopic Ligands in G Protein-Coupled Receptors.** *AAPS J* **2015,** *17* (5), 1080-95.

19. Feng, Z.; Ma, S.; Hu, G.; Xie, X. Q., **Allosteric Binding Site and Activation Mechanism of Class C G-Protein Coupled Receptors: Metabotropic Glutamate Receptor Family.** *AAPS J* **2015,** *17* (3), 737-53.

20. Ge, H.; Bian, Y.; He, X.; Xie, X. Q.; Wang, J., **Significantly different effects of tetrahydroberberrubine enantiomers on dopamine D1/D2 receptors revealed by experimental study and integrated in silico simulation.** *J Comput Aided Mol Des* **2019,** *33* (4), 447-459.

21. Gur, M.; Cheng, M. H.; Zomot, E.; Bahar, I., **Effect of Dimerization on the Dynamics of Neurotransmitter:Sodium Symporters.** *J Phys Chem B* **2017,** *121* (15), 3657-3666.

22. Gur, M.; Zomot, E.; Cheng, M. H.; Bahar, I., **Energy landscape of LeuT from molecular simulations.** *J Chem Phys* **2015,** *143* (24), 243134.

23. Hu, J.; Feng, Z.; Ma, S.; Zhang, Y.; Tong, Q.; Alqarni, M. H.; Gou, X.; Xie, X. Q., **Difference and Influence of Inactive and Active States of Cannabinoid Receptor Subtype CB2: From Conformation to Drug Discovery.** *J Chem Inf Model* **2016,** *56* (6), 1152-63.

24. Hu, Z.; Jing, Y.; Xue, Y.; Fan, P.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., **Analysis of substance use and its outcomes by machine learning: II. Derivation and prediction of the trajectory of substance use severity.** *Drug and Alcohol Dependence* **2019**, In press.

25. Ji, B.; Liu, S.; Xue, Y.; He, X.; Man, V. H.; Xie, X.-Q.; Wang, J**., Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation.** *Drugs R & D* **2019,** *19* (3), 297-305.

26. Jing, Y.; Hu, Z.; Fan, P.; Xue, Y.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., **Analysis of substance use and its outcomes by machine learning I. Childhood Evaluation of Liability to Substance Use Disorder.** *Drug And Alcohol Dependence* **2019**, In press.

27. Jun, I.; Cheng, M. H.; Sim, E.; Jung, J.; Suh, B. L.; Kim, Y.; Son, H.; Park, K.; Kim, C. H.; Yoon, J. H.; Whitcomb, D. C.; Bahar, I.; Lee, M. G., **Pore dilatation increases the bicarbonate permeability of CFTR, ANO1 and glycine receptor anion channels.** *J Physiol* **2016,** *594* (11), 2929-55.

28. Kaya, C.; Cheng, M. H.; Block, E. R.; Bartol, T. M.; Sejnowski, T. J.; Sorkin, A.; Faeder, J. R.; Bahar, I., **Heterogeneities in Axonal Structure and Transporter Distribution Lower Dopamine Reuptake Efficiency.** *eNeuro* **2018,** *5* (1).

29. Krieger, J.; Bahar, I.; Greger, I. H., **Structure, Dynamics, and Allosteric Potential of Ionotropic Glutamate Receptor N-Terminal Domains.** *Biophys J* **2015,** *109* (6), 1136-48.

30. Krieger, J.; Lee, J. Y.; Greger, I. H.; Bahar, I., **Activation and desensitization of ionotropic glutamate receptors by selectively triggering pre-existing motions.** *Neurosci Lett* **2019,** *700*, 22-29.

31. Lee, S.; Kong, S.; Xing, E. P., **A network-driven approach for genome-wide association mapping.** *Bioinformatics* **2016,** *32* (12), i164-i173.

32. Ma, S.; Cheng, M. H.; Guthrie, D. A.; Newman, A. H.; Bahar, I.; Sorkin, A., **Targeting of dopamine transporter to filopodia requires an outward-facing conformation of the transporter.** *Sci Rep* **2017,** *7* (1), 5399.

33. Man, V. H.; He, X.; Derreumaux, P.; Ji, B.; Xie, X. Q.; Nguyen, P. H.; Wang, J., **Effects of All-Atom Molecular Mechanics Force Fields on Amyloid Peptide Assembly: The Case of Abeta16-22 Dimer.** *J Chem Theory Comput* **2019,** *15* (2), 1440-1452.

34. Pei, F.; Li, H.; Liu, B.; Bahar, I., **Quantitative Systems Pharmacological Analysis of Drugs of Abuse Reveals the Pleiotropy of Their Targets and the Effector Role of mTORC1.** *Front Pharmacol* **2019,** *10*, 191.

35. Ponzoni, L.; Zhang, S.; Cheng, M. H.; Bahar, I., **Shared dynamics of LeuT superfamily members and allosteric differentiation by structural irregularities and multimerization.** *Philos Trans R Soc Lond B Biol Sci* **2018,** *373* (1749).

36. van Dijk, L.; Giladi, M.; Refaeli, B.; Hiller, R.; Cheng, M. H.; Bahar, I.; Khananshvili, D., **Key residues controlling bidirectional ion movements in Na(+)/Ca(2+) exchanger.** *Cell Calcium* **2018,** *76*, 10-22.

37. Wang, L.; Xie, X. Q., **Computational target fishing: what should chemogenomics researchers expect for the future of in silico drug design and discovery?** *Future Med Chem* **2014,** *6* (3), 247-9.

38. Wang, Y.; Lin, W.; Wu, N.; He, X.; Wang, J.; Feng, Z.; Xie, X. Q., **An insight into paracetamol and its metabolites using molecular docking and molecular dynamics simulation.** *J Mol Model* **2018,** *24* (9), 243.

39. Wang, Y. Q.; Lin, W. W.; Wu, N.; Wang, S. Y.; Chen, M. Z.; Lin, Z. H.; Xie, X. Q.; Feng, Z. W., **Structural insight into the serotonin (5-HT) receptor family by molecular docking, molecular dynamics simulation and systems pharmacology analysis.** *Acta Pharmacol Sin* **2019,** *40*, 1138–1156.

40. Wu, N.; Feng, Z.; He, X.; Kwon, W.; Wang, J.; Xie, X. Q., **Insight of Captagon Abuse by Chemogenomics Knowledgebase-guided Systems Pharmacology Target Mapping Analyses.** *Sci Rep* **2019,** *9* (1), 2268.

41. Wu, X.; Xie, S.; Wang, L.; Fan, P.; Ge, S.; Xie, X. Q.; Wu, W., **A computational strategy for finding novel targets and therapeutic compounds for opioid dependence.** *PLoS One* **2018,** *13* (11), e0207027.

42. Xie, X. Q.; Wang, L.; Liu, H.; Ouyang, Q.; Fang, C.; Su, W., **Chemogenomics knowledgebased polypharmacology analyses of drug abuse related G-protein coupled receptors and their ligands.** *Front Pharmacol* **2014,** *5*, 3.

43. Xu, X.; Ma, S.; Feng, Z.; Hu, G.; Wang, L.; Xie, X. Q., **Chemogenomics knowledgebase and systems pharmacology for hallucinogen target identification-Salvinorin A as a case study.** *J Mol Graph Model* **2016,** *70*, 284-295.

44. Zhang, S.; Jia, N.; Shao, P.; Tong, Q.; Xie, X. Q.; Bai, M., **Target-selective phototherapy using a ligand-based photosensitizer for type 2 cannabinoid receptor.** *Chem Biol* **2014,** *21* (3), 338-44.

45. Zhou, L.; Zhou, S.; Yang, P.; Tian, Y.; Feng, Z.; Xie, X. Q.; Liu, Y., **Targeted inhibition of the type 2 cannabinoid receptor is a novel approach to reduce renal fibrosis.** *Kidney Int* **2018,** *94* (4), 756-772.

46. Zomot, E.; Gur, M.; Bahar, I., **Microseconds simulations reveal a new sodium-binding site and the mechanism of sodium-coupled substrate uptake by LeuT.** *J Biol Chem* **2015,** *290* (1), 544-55.

II. DA-related health issues (i.e., asthma, cardiovascular problems) or DA-assoc. cardiotoxicity (also NIDA-funded) (13)

47. Anthonymuthu, T. S.; Kenny, E. M.; Shrivastava, I.; Tyurina, Y. Y.; Hier, Z. E.; Ting, H.-C.; Dar, H. H.; Tyurin, V. A.; Nesterova, A.; Amoscato, A. A.; Mikulska-Ruminska, K.; Rosenbaum, J. C.; Mao, G.; Zhao, J.; Conrad, M.; Kellum, J. A.; Wenzel, S. E.; VanDemark, A. P.; Bahar, I.; Kagan, V. E.; Bayır, H., **Empowerment of 15-lipoxygenase catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE.** *Journal of the American Chemical Society* **2018,** *140* (51), 17835-17839.

48. Bang, S.-J.; Wu, W. In ***Naive Bayes ensemble: A new approach to classifying unlabeled multi-class asthma subjects***, 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE: 2016; pp 460-465.

49. Feng, Z.; Pearce, L. V.; Xu, X.; Yang, X.; Yang, P.; Blumberg, P. M.; Xie, X. Q., **Structural insight into tetrameric hTRPV1 from homology modeling, molecular docking, molecular dynamics simulation, virtual screening, and bioassay validations.** *J Chem Inf Model* **2015,** *55* (3), 572-88.

50. Feng, Z.; Pearce, L. V.; Zhang, Y.; Xing, C.; Herold, B. K.; Ma, S.; Hu, Z.; Turcios, N. A.; Yang, P.; Tong, Q.; McCall, A. K.; Blumberg, P. M.; Xie, X. Q., **Multi-Functional Diarylurea Small Molecule Inhibitors of TRPV1 with Therapeutic Potential for Neuroinflammation.** *AAPS J* **2016,** *18* (4), 898-913.

51. Gao, Y.; Yang, P.; Shen, H.; Yu, H.; Song, X.; Zhang, L.; Zhang, P.; Cheng, H.; Xie, Z.; Hao, S.; Dong, F.; Ma, S.; Ji, Q.; Bartlow, P.; Ding, Y.; Wang, L.; Liu, H.; Li, Y.; Cheng, H.; Miao, W.; Yuan, W.; Yuan, Y.; Cheng, T.; Xie, X. Q., **Small-molecule inhibitors targeting INK4 protein p18(INK4C) enhance ex vivo expansion of haematopoietic stem cells.** *Nat Commun* **2015,** *6*, 6328.

52. Howrylak, J. A.; Moll, M.; Weiss, S. T.; Raby, B. A.; Wu, W.; Xing, E. P., **Gene expression profiling of asthma phenotypes demonstrates molecular signatures of atopy and asthma control.** *J Allergy Clin Immunol* **2016,** *137* (5), 1390-1397 e6.

53. Marchetti-Bowick, M.; Yu, Y.; Wu, W.; Xing, E. P., **A penalized regression model for the joint estimation of eQTL associations and gene network structure.** *The Annals of Applied Statistics* **2019,** *13* (1), 248-270.

54. Wu, W.; Bang, S.; Bleecker, E. R.; Castro, M.; Denlinger, L.; Erzurum, S. C.; Fahy, J. V.; Fitzpatrick, A. M.; Gaston, B. M.; Hastie, A. T.; Israel, E.; Jarjour, N. N.; Levy, B. D.; Mauger, D. T.; Meyers, D. A.; Moore, W. C.; Peters, M.; Phillips, B. R.; Phipatanakul, W.; Sorkness, R. L.; Wenzel, S. E., **Multiview Cluster Analysis Identifies Variable Corticosteroid Response Phenotypes in Severe Asthma.** *Am J Respir Crit Care Med* **2019,** *199* (11), 1358-1367.

55. Wu, W.; Bleecker, E.; Moore, W.; Busse, W. W.; Castro, M.; Chung, K. F.; Calhoun, W. J.; Erzurum, S.; Gaston, B.; Israel, E.; Curran-Everett, D.; Wenzel, S. E., **Unsupervised phenotyping of Severe Asthma Research Program participants using expanded lung data.** *J Allergy Clin Immunol* **2014,** *133* (5), 1280-8.

56. Xie, X. Q.; Yang, P.; Zhang, Y.; Zhang, P.; Wang, L.; Ding, Y.; Yang, M.; Tong, Q.; Cheng, H.; Ji, Q.; McGuire, T.; Yuan, W.; Cheng, T.; Gao, Y., **Discovery of novel INK4C small-molecule inhibitors to promote human and murine hematopoietic stem cell ex vivo expansion.** *Sci Rep* **2015,** *5*, 18115.

57. Xue, Y.; Feng, Z. W.; Li, X. Y.; Hu, Z. H.; Xu, Q.; Wang, Z.; Cheng, J. H.; Shi, H. T.; Wang, Q. B.; Wu, H. Y.; Xie, X. Q.; Lv, Q. Z., **The efficacy and safety of cilostazol as an alternative to aspirin in Chinese patients with aspirin intolerance after coronary stent implantation: a combined clinical study and computational system pharmacology analysis.** *Acta Pharmacol Sin* **2018,** *39* (2), 205-212.

58. Zhang, H.; Ma, S.; Feng, Z.; Wang, D.; Li, C.; Cao, Y.; Chen, X.; Liu, A.; Zhu, Z.; Zhang, J.; Zhang, G.; Chai, Y.; Wang, L.; Xie, X. Q., **Cardiovascular Disease Chemogenomics Knowledgebase-guided Target Identification and Drug Synergy Mechanism Study of an Herbal Formula.** *Sci Rep* **2016,** *6*, 33963.

59. Zhang, Y.; Wang, L.; Feng, Z.; Cheng, H.; McGuire, T. F.; Ding, Y.; Cheng, T.; Gao, Y.; Xie, X. Q., **StemCellCKB: An Integrated Stem Cell-Specific Chemogenomics KnowledgeBase for Target Identification and Systems-Pharmacology Research.** *J Chem Inf Model* **2016,** *56* (10), 1995-2004.

III. DA-associated neurological disorders (NDs) (16)

60. Adamik, J.; Silbermann, R.; Marino, S.; Sun, Q.; Anderson, J. L.; Zhou, D.; Xie, X. Q.; Roodman, G. D.; Galson, D. L., **XRK3F2 Inhibition of p62-ZZ Domain Signaling Rescues Myeloma-Induced GFI1-Driven Epigenetic Repression of the Runx2 Gene in Pre-osteoblasts to Overcome Differentiation Suppression.** *Front Endocrinol (Lausanne)* **2018,** *9*, 344.

61. Cha-Molstad, H.; Lee, S. H.; Kim, J. G.; Sung, K. W.; Hwang, J.; Shim, S. M.; Ganipisetti, S.; McGuire, T.; Mook-Jung, I.; Ciechanover, A.; Xie, X. Q.; Kim, B. Y.; Kwon, Y. T., **Regulation of autophagic proteolysis by the N-recognin SQSTM1/p62 of the N-end rule pathway.** *Autophagy* **2018,** *14* (2), 359-361.

62. Cha-Molstad, H.; Yu, J. E.; Feng, Z.; Lee, S. H.; Kim, J. G.; Yang, P.; Han, B.; Sung, K. W.; Yoo, Y. D.; Hwang, J.; McGuire, T.; Shim, S. M.; Song, H. D.; Ganipisetti, S.; Wang, N.; Jang, J. M.; Lee, M. J.; Kim, S. J.; Lee, K. H.; Hong, J. T.; Ciechanover, A.; Mook-Jung, I.; Kim, K. P.; Xie, X. Q.; Kwon, Y. T.; Kim, B. Y., **p62/SQSTM1/Sequestosome-1 is an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis.** *Nat Commun* **2017,** *8* (1), 102.

63. Fan, P.; Wang, N.; Wang, L.; Xie, X. Q., Autophagy And Apoptosis Specifc Knowledgebases-Guided Systems Pharmacology Drug Research. *Curr Cancer Drug Targets* **2019**.

64. Hu, Z.; Wang, L.; Ma, S.; Kirisci, L.; Feng, Z.; Xue, Y.; Klunk, W. E.; Kamboh, M. I.; Sweet, R. A.; Becker, J.; Lv, Q.; Lopez, O. L.; Xie, X. Q., **Synergism of antihypertensives and cholinesterase inhibitors in Alzheimer's disease.** *Alzheimers Dement (N Y)* **2018,** *4*, 542-555.

65. Lee, J. Y.; Feng, Z.; Xie, X. Q.; Bahar, I., **Allosteric Modulation of Intact gamma-Secretase Structural Dynamics.** *Biophys J* **2017,** *113* (12), 2634-2649.

66. Lee, S.; Lozano, A.; Kambadur, P.; Xing, E. P., **An Efficient Nonlinear Regression Approach for Genome-wide Detection of Marginal and Interacting Genetic Variations.** *J Comput Biol* **2016,** *23* (5), 372-89.

67. Lee, S.; Wang, H.; Xing, E. P., **Backward genotype-transcript-phenotype association mapping.** *Methods* **2017,** *129*, 18-23.

68. Liu, H.; Wang, L.; Lv, M.; Pei, R.; Li, P.; Pei, Z.; Wang, Y.; Su, W.; Xie, X. Q., **AlzPlatform: an Alzheimer's disease domain-specific chemogenomics knowledgebase for polypharmacology and target identification research.** *J Chem Inf Model* **2014,** *54* (4), 1050-60.

69. Liu, H.; Wang, L.; Su, W.; Xie, X. Q., **Advances in recent patent and clinical trial drug development for Alzheimer's disease.** *Pharm Pat Anal* **2014,** *3* (4), 429-47.

70. Ma, S.; Attarwala, I. Y.; Xie, X. Q., **SQSTM1/p62: A Potential Target for Neurodegenerative Disease.** *ACS Chem Neurosci* **2019,** *10* (5), 2094-2114.

71. Pei, F.; Li, H.; Henderson, M. J.; Titus, S. A.; Jadhav, A.; Simeonov, A.; Cobanoglu, M. C.; Mousavi, S. H.; Shun, T.; McDermott, L.; Iyer, P.; Fioravanti, M.; Carlisle, D.; Friedlander, R. M.; Bahar, I.; Taylor, D. L.; Lezon, T. R.; Stern, A. M.; Schurdak, M. E., **Connecting Neuronal Cell Protective Pathways and Drug Combinations in a Huntington's Disease Model through the Application of Quantitative Systems Pharmacology.** *Sci Rep* **2017,** *7* (1), 17803.

72. Wang, H.; Aragam, B.; Xing, E. P., **Variable Selection in Heterogeneous Datasets: A Truncated-rank Sparse Linear Mixed Model with Applications to Genome-wide Association Studies.** *Proceedings (IEEE Int Conf Bioinformatics Biomed)* **2017,** *2017*, 431-438.

73. Wang, K. Z. Q.; Steer, E.; Otero, P. A.; Bateman, N. W.; Cheng, M. H.; Scott, A. L.; Wu, C.; Bahar, I.; Shih, Y. T.; Hsueh, Y. P.; Chu, C. T., **PINK1 Interacts with VCP/p97 and Activates PKA to Promote NSFL1C/p47 Phosphorylation and Dendritic Arborization in Neurons.** *eNeuro* **2018,** *5* (6).

74. Wang, L.; Ma, S.; Hu, Z.; McGuire, T. F.; Xie, X.-Q., **Chemogenomics Systems Pharmacology Mapping of Potential Drug Targets for Treatment of Traumatic Brain Injury.** *J Neurotrauma* **2019,** *36* (4), 565-575.

75. Yoo, Y. D.; Mun, S. R.; Ji, C. H.; Sung, K. W.; Kang, K. Y.; Heo, A. J.; Lee, S. H.; An, J. Y.; Hwang, J.; Xie, X. Q.; Ciechanover, A.; Kim, B. Y.; Kwon, Y. T., **N-terminal arginylation generates a bimodal degron that modulates autophagic proteolysis.** *Proc Natl Acad Sci U S A* **2018,** *115* (12), E2716-E2724.

IV. Development of methods and software, as we proposed in the 1st cycle proposal (52)

76. Al-Shedivat, M.; Wilson, A. G.; Saatchi, Y.; Hu, Z.; Xing, E. P., **Learning Scalable Deep Kernels with Recurrent Structure.** *J Mach Learn Res* **2017,** *18* (1), 2850-2886.

77. Bahar, I.; Cheng, M. H.; Lee, J. Y.; Kaya, C.; Zhang, S., **Structure-Encoded Global Motions and Their Role in Mediating Protein-Substrate Interactions.** *Biophys J* **2015,** *109* (6), 1101-9.

78. Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I., **Evol and ProDy for bridging protein sequence evolution and structural dynamics.** *Bioinformatics* **2014,** *30* (18), 2681-3.

79. Bian, Y.; Xie, X. S., **Computational Fragment-Based Drug Design: Current Trends, Strategies, and Applications.** *AAPS J* **2018,** *20* (3), 59.

80. Chang, X.; Yang, Y.; Xing, E.; Yu, Y. In ***Complex event detection using semantic saliency and nearly-isotonic SVM***, International Conference on Machine Learning, 2015; pp 1348-1357.

81. Cobanoglu, M. C.; Oltvai, Z. N.; Taylor, D. L.; Bahar, I., **BalestraWeb: efficient online evaluation of drug-target interactions.** *Bioinformatics* **2015,** *31* (1), 131-3.

82. Dan, C.; Leqi, L.; Aragam, B.; Ravikumar, P. K.; Xing, E. P. In ***The sample complexity of semi-supervised learning with nonparametric mixture models***, Advances in Neural Information Processing Systems, 2018; pp 9321-9332.

83. Haliloglu, T.; Bahar, I., **Adaptability of protein structures to enable functional interactions and evolutionary implications.** *Curr Opin Struct Biol* **2015,** *35*, 17-23.

84. He, X.; Man, V. H.; Ji, B.; Xie, X. Q.; Wang, J., **Calculate protein-ligand binding affinities with the extended linear interaction energy method: application on the Cathepsin S set in the D3R Grand Challenge 3.** *J Comput Aided Mol Des* **2019,** *33* (1), 105-117.

85. Jing, Y.; Bian, Y.; Hu, Z.; Wang, L.; Xie, X. Q., **Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era.** *AAPS J* **2018,** *20* (3), 58.

86. Kolar, M.; Liu, H.; Xing, E. P., **Graph Estimation From Multi-Attribute Data.** *J Mach Learn Res* **2014,** *15* (May), 1713-1750.

87. Kurkcuoglu, Z.; Bahar, I.; Doruker, P**., ClustENM: ENM-Based Sampling of Essential Conformational Space at Full Atomic Resolution.** *J Chem Theory Comput* **2016,** *12* (9), 4549-62.

88. Lee, J. Y.; Krieger, J.; Herguedas, B.; Garcia-Nafria, J.; Dutta, A.; Shaikh, S. A.; Greger, I. H.; Bahar, I., **Druggability Simulations and X-Ray Crystallography Reveal a Ligand-Binding Site in the GluA3 AMPA Receptor N-Terminal Domain.** *Structure* **2019,** *27* (2), 241-252 e3.

89. Lee, S.; Gornitz, N.; Xing, E. P.; Heckerman, D.; Lippert, C., Ensembles of Lasso Screening Rules. *IEEE Trans Pattern Anal Mach Intell* **2018,** *40* (12), 2841-2852.

90. Lengerich, B. J.; Aragam, B.; Xing, E. P., **Personalized regression enables sample-specific pan-cancer analysis.** *Bioinformatics* **2018,** *34* (13), i178-i186.

91. Li, H.; Chang, Y. Y.; Lee, J. Y.; Bahar, I.; Yang, L. W., **DynOmics: dynamics of structural proteome and beyond.** *Nucleic Acids Res* **2017,** *45* (W1), W374-W380.

92. Liu, B.; Gyori, B. M.; Thiagarajan, P., **Statistical Model Checking-Based Analysis of Biological Networks.** In *Automated Reasoning for Systems Biology and Medicine*, Springer: 2019; pp 63-92.

93. Man, V. H.; Li, M. S.; Wang, J.; Derreumaux, P.; Nguyen, P. H., **Interaction mechanism between the focused ultrasound and lipid membrane at the molecular level.** *J Chem Phys* **2019,** *150* (21), 215101.

94. Man, V. H.; Li, M. S.; Wang, J.; Derreumaux, P.; Nguyen, P. H., **Nonequilibrium atomistic molecular dynamics simulation of tubular nanomotor propelled by bubble propulsion.** *J Chem Phys* **2019,** *151* (2), 024103.

95. Man, V. H.; Truong, P. M.; Li, M. S.; Wang, J.; Van-Oanh, N. T.; Derreumaux, P.; Nguyen, P. H., **Molecular Mechanism of the Cell Membrane Pore Formation Induced by Bubble Stable Cavitation.** *J Phys Chem B* **2019,** *123* (1), 71-78.

96. Marchetti-Bowick, M.; Yin, J.; Howrylak, J. A.; Xing, E. P., **A time-varying group sparse additive model for genome-wide association studies of dynamic complex traits.** *Bioinformatics* **2016,** *32* (19), 2903-10.

97. Myint, K. Z.; Xie, X. Q., **Ligand biological activity predictions using fingerprint-based artificial neural networks (FANN-QSAR).** *Methods Mol Biol* **2015,** *1260*, 149-64.

98. Neiswanger, W.; Xing, E. In ***Post-inference prior swapping***, Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR. org: 2017; pp 2594-2602.

99. Qin, L.; Zhang, Z.; Zhao, H.; Hu, Z.; Xing, E. P., **Adversarial connective-exploiting networks for implicit discourse relation classification.** *arXiv preprint arXiv:1704.00217* **2017**.

100. Sachan, M.; Dubey, K.; Xing, E. In ***From textbooks to knowledge: A case study in harvesting axiomatic knowledge from textbooks to solve geometry problems***, Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017; pp 773-784.

101. Sachan, M.; Xing, E. In ***Self-training for jointly learning to ask and answer questions***, Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), 2018; pp 629-640.

102. Sauerwald, N.; Zhang, S.; Kingsford, C.; Bahar, I., **Chromosomal dynamics predicted by an elastic network model explains genome-wide accessibility and long-range couplings.** *Nucleic Acids Res* **2017,** *45* (7), 3663-3673.

103. Stern, A. M.; Schurdak, M. E.; Bahar, I.; Berg, J. M.; Taylor, D. L., **A Perspective on Implementing a Quantitative Systems Pharmacology Platform for Drug Discovery and the Advancement of Personalized Medicine.** *J Biomol Screen* **2016,** *21* (6), 521-34.

104. Wang, H.; Lengerich, B. J.; Aragam, B.; Xing, E. P., **Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data.** *Bioinformatics* **2019,** *35* (7), 1181-1187.

105. Wang, H.; Liu, X.; Tao, Y.; Ye, W.; Jin, Q.; Cohen, W. W.; Xing, E. P., **Automatic Human-like Mining and Constructing Reliable Genetic Association Database with Deep Reinforcement Learning.** *Pac Symp Biocomput* **2019,** *24*, 112-123.

106. Wang, H.; Liu, X.; Xiao, Y.; Xu, M.; Xing, E. P., **Multiplex confounding factor correction for genomic association mapping with squared sparse linear mixed model.** *Methods* **2018,** *145*, 33-40.

107. Wang, H.; Wu, Z.; Xing, E. P., **Removing Confounding Factors Associated Weights in Deep Neural Networks Improves the Prediction Accuracy for Healthcare Applications.** *Pac Symp Biocomput* **2019,** *24*, 54-65.

108. Wang, J.; Ge, Y.; Xie, X. Q., **Development and Testing of Druglike Screening Libraries.** *J Chem Inf Model* **2019,** *59* (1), 53-65.

109. Wang, L.; Xie, X. Q., **Cancer genomics: opportunities for medicinal chemistry?** *Future Med Chem* **2016,** *8* (4), 357-9.

110. Wang, N.; Wang, L.; Xie, X. Q., **ProSelection: A Novel Algorithm to Select Proper Protein Structure Subsets for in Silico Target Identification and Drug Discovery Research.** *J Chem Inf Model* **2017,** *57* (11), 2686-2698.

111. Xiaojun, C.; Yao-Liang, Y.; Yi, Y.; Xing, E. P., **Semantic Pooling for Complex Event Analysis in Untrimmed Videos.** *IEEE Trans Pattern Anal Mach Intell* **2017,** *39* (8), 1617-1632.

112. Xie, P.; Deng, Y.; Zhou, Y.; Kumar, A.; Yu, Y.; Zou, J.; Xing, E. P. In ***Learning latent space models with angular constraints***, Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR. org: 2017; pp 3799-3810.

113. Xie, P.; Poczos, B.; Xing, E. P. In ***Near-Orthogonality Regularization in Kernel Methods***, UAI, 2017; p 6.

114. Xie, P.; Salakhutdinov, R.; Mou, L.; Xing, E. P. In ***Deep determinantal point process for large-scale multi-label classification***, Proceedings of the IEEE International Conference on Computer Vision, 2017; pp 473-482.

115. Xie, P.; Singh, A.; Xing, E. P. In ***Uncorrelation and evenness: a new diversity-promoting regularizer***, Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR. org: 2017; pp 3811-3820.

116. Xie, P.; Wu, W.; Zhu, Y.; Xing, E. P., **Orthogonality-promoting distance metric learning: Convex relaxation and theoretical analysis.** *arXiv preprint arXiv:1802.06014* **2018**.

117. Xie, P.; Xing, E. In ***A constituent-centric neural architecture for reading comprehension***, Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2017; pp 1405-1414.

118. Xie, P.; Xing, E. In ***A neural architecture for automated icd coding*****,** Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018; pp 1066-1076.

119. Xie, P.; Zhang, H.; Zhu, Y.; Xing, E. In ***Nonoverlap-promoting variable selection***, International Conference on Machine Learning, 2018; pp 5413-5422.

120. Xing, E. P.; Curtis, R. E.; Schoenherr, G.; Lee, S.; Yin, J.; Puniyani, K.; Wu, W.; Kinnaird, P., **GWAS in a box: statistical and visual analytics of structured associations via GenAMap.** *PLoS One* **2014,** *9* (6), e97524.

121. Xu, M.; Chai, X.; Muthakana, H.; Liang, X.; Yang, G.; Zeev-Ben-Mordehai, T.; Xing, E. P., **Deep learning-based subdivision approach for large scale macromolecules structure recovery from electron cryo tomograms.** *Bioinformatics* **2017,** *33* (14), i13-i22.

122. Zhang, H.; Zheng, Z.; Xu, S.; Dai, W.; Ho, Q.; Liang, X.; Hu, Z.; Wei, J.; Xie, P.; Xing, E. P. In ***Poseidon: An efficient communication architecture for distributed deep learning on {GPU} clusters***, 2017 {USENIX} Annual Technical Conference ({USENIX}{ATC} 17), 2017; pp 181-193.

123. Zhang, S.; Li, H.; Krieger, J. M.; Bahar, I., **Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity.** *Mol Biol Evol* **2019,** *36* (9), 2053–2068.

124. Zheng, X.; Yu, Y.; Xing, E. P. In ***Linear time samplers for supervised topic models using compositional proposals***, Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM: 2015; pp 1523-1532.

125. Zhou, Y.; Yuan, K.; Yu, Y.; Ni, X.; Xie, P.; Xing, E. P.; Xu, S., **Inference of multiple-wave population admixture by modeling decay of linkage disequilibrium with polynomial functions.** *Heredity (Edinb)* **2017,** *118* (5), 503-510.

126. Taylor, D. L.; Gough, A.; Schurdak, M. E.; Vernetti, L.; Chennubhotla, C. S.; Lefever, D.; Pei, F.; Faeder, J. R.; Lezon, T. R.; Stern, A. M.; Bahar, I., **Harnessing Human Microphysiology Systems as Key Experimental Models for Quantitative Systems Pharmacology.** Handbook of Experimental Pharmacology: 2019; pp 1-41.

127. Lee, J.; Krieger, J.; Li, H.; Bahar, I., **Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations.** *Protein Science* **2019**, Accepted.

V. Collaborative studies with researchers using our resources (16)

128. Cai, Z.; Ouyang, Q.; Zeng, D.; Nguyen, K. N.; Modi, J.; Wang, L.; White, A. G.; Rogers, B. E.; Xie, X. Q.; Anderson, C. J., **64Cu-labeled somatostatin analogues conjugated with cross-bridged phosphonate-based chelators via strain-promoted click chemistry for PET imaging: in silico through in vivo studies.** *J Med Chem* **2014,** *57* (14), 6019-29.

129. Chen, S.; Feng, Z.; Wang, Y.; Ma, S.; Hu, Z.; Yang, P.; Chai, Y.; Xie, X., **Discovery of Novel Ligands for TNF-alpha and TNF Receptor-1 through Structure-Based Virtual Screening and Biological Assay.** *J Chem Inf Model* **2017,** *57* (5), 1101-1111.

130. Chen, X.; Cao, Y.; Zhang, H.; Zhu, Z.; Liu, M.; Liu, H.; Ding, X.; Hong, Z.; Li, W.; Lv, D.; Wang, L.; Zhuo, X.; Zhang, J.; Xie, X. Q.; Chai, Y., **Comparative normal/failing rat myocardium cell membrane chromatographic analysis system for screening specific components that counteract doxorubicin-induced heart failure from Acontium carmichaeli.** *Anal Chem* **2014,** *86* (10), 4748-57.

131. Feng, Z.; Kochanek, S.; Close, D.; Wang, L.; Srinivasan, A.; Almehizia, A. A.; Iyer, P.; Xie, X. Q.; Johnston, P. A.; Gold, B., **Design and activity of AP endonuclease-1 inhibitors.** *J Chem Biol* **2015,** *8* (3), 79-93.

132. Hu, J.; Hu, Z.; Zhang, Y.; Gou, X.; Mu, Y.; Wang, L.; Xie, X. Q., **Metal binding mediated conformational change of XPA protein:a potential cytotoxic mechanism of nickel in the nucleotide excision repair.** *J Mol Model* **2016,** *22* (7), 156.

133. Keskin, O.; Dyson, H. J.; Bahar, I., **Biomolecular Systems Interactions, Dynamics, and Allostery: Reflections and New Directions.** *Biophys J* **2015,** *109* (6), E01-2.

134. Liu, B.; Liu, Q.; Yang, L.; Palaniappan, S. K.; Bahar, I.; Thiagarajan, P. S.; Ding, J. L., **Innate immune memory and homeostasis may be conferred through crosstalk between the TLR3 and TLR7 pathways.** *Sci Signal* **2016,** *9* (436), ra70.

135. Mikulska-Ruminska, K.; Shrivastava, I.; Krieger, J.; Zhang, S.; Li, H.; Bayir, H.; Wenzel, S. E.; VanDemark, A. P.; Kagan, V. E.; Bahar, I., **Characterization of Differential Dynamics, Specificity, and Allostery of Lipoxygenase Family Members.** *J Chem Inf Model* **2019,** *59* (5), 2496-2508.

136. Ouyang, Q.; Wang, L.; Mu, Y.; Xie, X. Q., **Modeling skin sensitization potential of mechanistically hard-to-be-classified aniline and phenol compounds with quantum mechanistic properties.** *BMC Pharmacol Toxicol* **2014,** *15*, 76.

137. Stoyanovsky, D. A.; Tyurina, Y. Y.; Shrivastava, I.; Bahar, I.; Tyurin, V. A.; Protchenko, O.; Jadhav, S.; Bolevich, S. B.; Kozlov, A. V.; Vladimirov, Y. A.; Shvedova, A. A.; Philpott, C. C.; Bayir, H.; Kagan, V. E., **Iron catalysis of lipid peroxidation in ferroptosis: Regulated enzymatic or random free radical reaction?** *Free Radic Biol Med* **2019,** *133*, 153-161.

138. Wenzel, S. E.; Tyurina, Y. Y.; Zhao, J.; St Croix, C. M.; Dar, H. H.; Mao, G.; Tyurin, V. A.; Anthonymuthu, T. S.; Kapralov, A. A.; Amoscato, A. A.; Mikulska-Ruminska, K.; Shrivastava, I. H.; Kenny, E. M.; Yang, Q.; Rosenbaum, J. C.; Sparvero, L. J.; Emlet, D. R.; Wen, X.; Minami, Y.; Qu, F.; Watkins, S. C.; Holman, T. R.; VanDemark, A. P.; Kellum, J. A.; Bahar, I.; Bayir, H.; Kagan, V. E., **PEBP1 Wardens Ferroptosis by Enabling Lipoxygenase Generation of Lipid Death Signals.** *Cell* **2017,** *171* (3), 628-641 e26.

139. Wodak, S. J.; Paci, E.; Dokholyan, N. V.; Berezovsky, I. N.; Horovitz, A.; Li, J.; Hilser, V. J.; Bahar, I.; Karanicolas, J.; Stock, G.; Hamm, P.; Stote, R. H.; Eberhardt, J.; Chebaro, Y.; Dejaegere, A.; Cecchini, M.; Changeux, J. P.; Bolhuis, P. G.; Vreede, J.; Faccioli, P.; Orioli, S.; Ravasio, R.; Yan, L.; Brito, C.; Wyart, M.; Gkeka, P.; Rivalta, I.; Palermo, G.; McCammon, J. A.; Panecka-Hofman, J.; Wade, R. C.; Di Pizio, A.; Niv, M. Y.; Nussinov, R.; Tsai, C. J.; Jang, H.; Padhorny, D.; Kozakov, D.; McLeish, T., **Allostery in Its Many Disguises: From Theory to Applications.** *Structure* **2019,** *27* (4), 566-578.

140. Yonjung Kim; Ikhyun Jun; Dong Hoon Shin; Jihoon G. Yoon; Jinsei Jung; Hyun Woo Park; Mary Hongying Cheng; Ivet Bahar; David C. Whitcomb; Lee, M. G., **Regulation of CFTR Bicarbonate Channel Activity by WNK1: Implications for Pancreatitis and CFTR-related disorders.** *Cellular and Molecular Gastroenterology and Hepatology* **2019**, In press.

141. Zhang, B.; Li, F.; Chen, Z.; Shrivastava, I. H.; Gasanoff, E. S.; Dagda, R. K**., Naja mossambica mossambica Cobra Cardiotoxin Targets Mitochondria to Disrupt Mitochondrial Membrane Structure and Function.** *Toxins (Basel)* **2019,** *11* (3).

142. Zhou, Z.; Feng, Z.; Hu, D.; Yang, P.; Gur, M.; Bahar, I.; Cristofanilli, M.; Gradishar, W. J.; Xie, X. Q.; Wan, Y., **A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy.** *EBioMedicine* **2019,** *44*, 98-111.

143. Zhou, Z.; Luo, A.; Shrivastava, I.; He, M.; Huang, Y.; Bahar, I.; Liu, Z.; Wan, Y., **Regulation of XIAP Turnover Reveals a Role for USP11 in Promotion of Tumorigenesis.** *EBioMedicine* **2017,** *15*, 48-61.

Website(s) or Other Internet Site(s)

The official website of CDAR (**http://www.cdarcenter.org**) has been operating well for 5 years, providing a web portal for fully accessing all of our computational tools and databases. This website

illustrates our personnel, research, technologies, outreach, publications, as well as meetings and news information of CDAR.

The AdminCore has also recently begun facilitating the implementation, dissemination, and broad use of CDAR resources via a **new Platform for Abused-Drugs and Neurological Diseases Association (PANDA)** (computational methods, software, application programming interfaces (APIs), databases(DBs), and GPU cloud servers) to benefit the FRPs, P/FPs, and the broader DAR community. PANDA (**http://www.cdarcenter.org/panda/**) will be a major resource that will integrate our tools facilitate the efficient usage of CDAR data and tools by our collaborators and the broader DA research (**DAR**) community. The overall structure of the PANDA and the roadmap of the task-driven webtool are illustrated in **Fig. 1**.

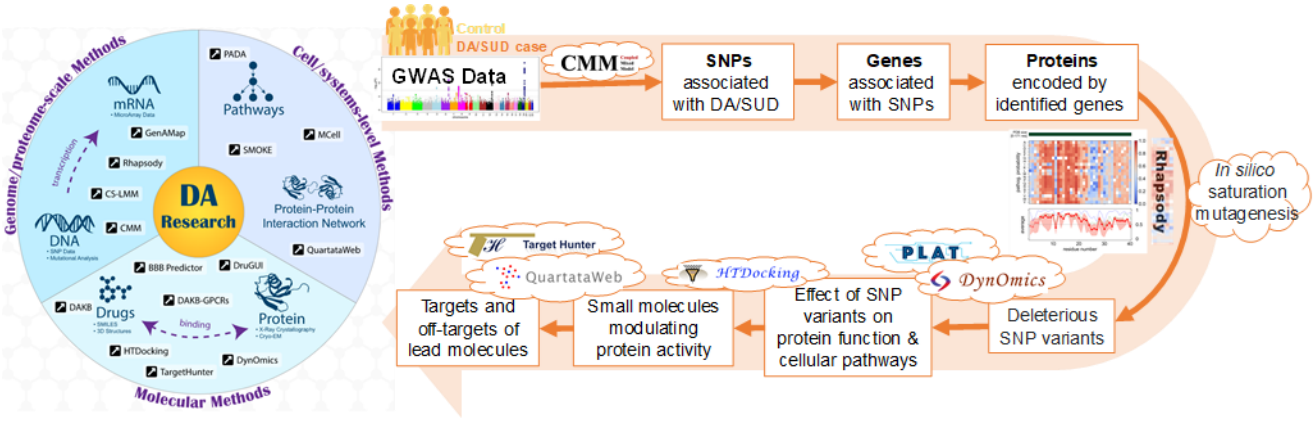


Fig 1. PANDA roadmap for computational analysis of DA/SUD clinical data at molecular-, genetic-, cellular-, and systems-levels and using CDAR tools and knowledgebases. The wheel diagram (*left*), from the PANDA website (http://www.cdarcenter.org/panda/) will be an interactive portal that provides access to the tools and databases developed by CDAR, along with task-driven protocols. The schematic (*right*) illustrates a potential protocol for a DA/SUD researcher to utilize CDAR tools. Starting from GWAS data, advanced gene analysis tools developed by Core C will identify SNPs associated with DA/SUD; proteins encoded by genes associated with the identified SNPs will be analyzed to determine their pathogenicity and the effects of genetic variants using tool RHAPSODY developed by Core B. Next, the pathogenic proteins are subjected to structure-based analyses of their functional dynamics to identify target sites, and high-throughput virtual screening of small molecules yields lead compounds that bind those sites using Core A tools. Promising drug candidates can be further assessed for potential off-target interactions using the integrated TargetHunter (Core A) and Quartataweb (Core B) servers.

Technologies or techniques

In the past years, we have developed a set of algorithms/tools/software and knowledgebases for DAR and rational drug design for broad-spectrum abused-/neurological drug targets. As a national resource for DAR and computer-aided drug discovery, our technologies have been extensively used by our peers, as summarized in Tables 6 – 8, below.

Core A Software (Table 6)

Technology	Description	URL
	Domain-specific Databases/Knowledgebases as National Resources	
CBID	A user-friendly web-interfaced cannabinoid molecular information database repository[64] with the integrated QSAR tools, e.g.,Fingerprint-based Artificial Neural Networks QSAR,[4] Fragment-based QSAR,[3] PharmShape-based QSAR,[89] 3D QSAR[65]	www.CBLigand.org/CBID
DAKB	Chemogenomics knowledgebase (**KB**) for drug abuse (**DA**)[14]	www.CBLigand.org/DAKB

AlzPlatform	Chemogenomics KB for Alzheimer's disease[2]	www.CBLigand.org/AD
Hallucinogen	Chemogenomics KB for Hallucinogen research[20]	www.CBLigand.org/hallucinogen
CVD	Chemogenomics KB for cardiovascular diseases[36]	www.cbligand.org/CVD
TBI	Chemogenomics KB for traumatic brain injury[46]	www.CBLigand.org/TBI
DLSL	Drug-like screening databases for general drug targets[55]	mulan.pharmacy.pitt.edu/database
Software & Web-Based Toolkits Accessible Online		
TargetHunter	Target/Off-target prediction for a small molecule(s)[6]	www.CBLigand.org/TargetHunter
HTDocking	High throughput docking for virtual screening[2]	www.CBLigand.org/HTDocking
LiCABEDS	Ligand Classifier of Adaptively Boosting Ensemble Decision Stumps (LiCABEDS), based on machine learning algorithm for compound classification prediction[5, 76]	www.CBLigand.org/LiCABEDS
PAINS-Remover	A Computer-Aided Drug Design (CADD) tool to identify the false positives in experimental high-throughput screening data	www.CBLigand.org/FPR
BBB	A CADD tool to predict the permeability of a chemical compound crossing blood brain barrier	www.CBligand.org/BBB
DAKB-GPCRs	An online GPCRs chemogenomics knowledgebase for DA research, implemented with our established chemogenomics system pharmacology (CSP-Target) mapping algorithms/tools for data visualization and analyses[1]	www.cbligand.org/dakb-gpcrs
Mol-Prop	GPU-accelerated molecular property calculation[71]	www.CBLigand.org/gpu
MMFFT	A user-friendly web toolkit for generating molecular mechanical force field (MMFF) models for arbitrary chemicals	mulan.pharmacy.pitt.edu/mmfft
re-Affinity	A software tool to re-rank docking poses using the MM-PB/GBSA scoring functions[52, 74, 80, 83, 90]	mulan.pharmacy.pitt.edu

* Newly launched applications are colored in red, those for download only are colored in green. **NA/NC**: not applied or not counted

Core B Software (Table 7)

Resources	Description
ProDy	an open-source Python package for protein structural dynamics; fast, flexible and powerful file parsers and customizable atom-selections for structural analysis. Its capabilities have recently been advanced (see below) http://prody.csb.pitt.edu/
ANM server	Analysis and visualization of the equilibrium motions of proteins deposited in the Protein Data Bank (PDB) http://anm.csb.pitt.edu/cgi-bin/anm2/anm2.cgi
coMD module	a hybrid method implemented in *ProDy*,[141] with ANM-predicted collective modes accelerating MD (coMD) simulations, extended[41] to map the energy landscape of LeuT-fold neurotransmitter transporters
Evol	a new *ProDy* module for bridging between sequence evolution and structural dynamics
iGNM DB	A database that provides access to the dynamics of 95% of structures available in the PDB http://ignm.ccbb.pitt.edu/ and http://gnm.csb.pitt.edu/
DynOmics	a portal to leverage structural proteomics data by evaluating structural dynamics (molecules, complexes and assemblies) in the context of their physiological environment (> 6,000 unique users since its inception in 2015)
BalestraWeb	Server for searching DrugBank and predicting new drug-target associations http://balestra.csb.pitt.edu/
ClustENM	A new algorithm for ANM-based sampling of essential conformational space at full atomic resolution
SMOKE	A statistical model checking tool for estimating unknown parameters of systems biology models
SignDy	A novel methodology for determining generic and specific aspects or family and subfamily functional dynamics
Pharmmaker	A novel tool for building pharmacophoric models, interoperating with DRUGUI and Pharmit for virtual screening
DruGUI	A VMD plugin designed for setup and analysis of simulations containing small organic molecules (probes) for druggability assessment http://prody.csb.pitt.edu/drugui/
Rhapsody	pathogenicity prediction of missense variants by taking structural dynamics into considerations (version 1) http://rhapsody.csb.pitt.edu/

Quartata	in silico chemogenomics methodology and server for linking drugs/chemicals, targets, pathways and GO annotations http://quartata.csb.pitt.edu

Core C Software (Table 8)

Resources	Description
GenA Map	GenAMap: A Visual Analytics Software Platform for eQTL and GWAS Analysis https://github.com/blengerich/GenAMap
Precision Lasso	Precision Lasso: Accounting for Correlations and Linear Dependencies in High-Dimensional Genomic Data. https://github.com/HaohanWang/thePrecisionLasso
CS-LMM	Constrained Sparse Linear Mixed Model: Discovering Weaker Genetic Associations with Validated Association https://github.com/HaohanWang/CS-LMM
DMM	Deep Mixed Model: Marginal Epistasis Detection and Population Stratification Correction in Genome-Wide Association Studies https://github.com/HaohanWang/DMM
MKKC	MKKC: An R-package For Multiple Kernel K-means Clustering. https://seojinbang.github.io/MKKC/

Inventions, Patent Applications, and/or Licenses

Number of patents: **2** (See the report from each individual core)
1 Joint disclosure filed

Data/Resource Sharing Plan

The University of Pittsburgh has a common policy on intellectual property. The Center and PIs have successfully worked and collaborated together for the past five years (the current funding cycle) and produced significant progress and achievements. Should IP or other issues become a concern, the University has a committee to address such problems.

We anticipate that our work will result in generation of large, information-rich datasets. Dissemination of findings and sharing of knowledge and copyright protected information will be available through the CDAR Network over the Internet. Established research resources including hardware and software will be shared with others who have interests in applying computational technologies available to the CDAR Center to their research projects. Computational protocols and data, once they are published, will also be shared with other researchers by following the NIH data sharing policy (http://grants.nih.gov/grants/policy/data_sharing). In addition, our data/resources sharing plan includes:

i) The sharing of documentation will be done through the CDAR Center web publishing mechanism and through the Center web server. For example, CDAR center will implement, establish and maintain new computing Platform for Abused Drugs and Neurological Disorders Associated with DA (**PANDA**) that will facilitate the efficient usage of CDAR data and computing resources/protocols by our collaborators and the broader DA research (**DAR**) community.

ii) The new computational algorithms and the data-mining programs will be available for academic users under a Materials Transfer Agreement and a data-sharing agreement defined by the participating Universities.

iii) New scaffold chemical probes or leads identified for drug abuse and neurological disease neurotherapy and associated 3D target protein/enzyme structures, as well as genomics data, will be deposited into the CDAR cloud server system and hyperlinked with other online databases for public access.

CORE A (PI: Dr. Xiang-Qun Xie)

- **AIMS**
- **ACCOMPLISHMENTS**
- **RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST**
- **PLAN FOR NEXT PERIOD**
- **PRODUCTS**
  - **PUBLICATIONS**
  - **WEBSITE(S) AND OTHER INTERNET SITE(S)**
  - **TECHNOLOGIES OR TECHNIQUES**
- **BIBLIOGRAPHY**

CORE A

COMPUTATIONAL CHEMOGENOMICS FOR DRUG ABUSE (CC4DA)

EXECUTIVE SUMMARY

Xiang-Qun Xie, MD PhD

MAJOR GOALS/AIMS OF THE PROJECT

Aim 1*:* To enable data sharing and processing among scientists in the DA and related scientific communities by our established chemogenomics knowledgebase for drug abuse (DA-KB)

Aim 2*:* To integrate and further advance our algorithms/tools for predicting DA targets, pathways, underlying mechanisms and potential polypharmacological effects relevant to polydrug addiction and DAR

Aim 3*:* To implement cloud sourcing and computing services to facilitate and accelerate computational chemogenomics studies as well as *in silico* medication design and discovery for DA therapeutics

ACCOMPLISHMENTS

In the first funding term (8/1/2014 – 7/31/2019, currently in NCE), Core A team has worked closely with **Cores B** and **C,** as well as the FRP and P/FP PIs, and made significant achievements towards the proposed research aims listed above. The major achievements for Core A in the first term include: (i) **58** peer-reviewed publications with 25 directly related to DA and 16 related to DA-associated health issues or DA-associated NDs, **23** of which were published after Oct. 2018, and most (34 papers) were achieved through collaborations; (ii) construction of **8** DA/ND-related databases; (iii) development and advancement of **15** innovative computational tools with online accessibility; (iv) support of **10** NIH-funded research projects (FRPs) and **3** pilot/feasibility projects (P/FPs) by Core A (http://www.cdarcenter.org); and (v) development of a new course, Pharmacometrics & Systems Pharmacology (**PSP**), for training the next generation of DA researchers (currently 5 PhD and 9 MS students) and organization of **2** training workshops.

The following is the summary of Core Technology Innovation and Research Accomplishments

1. **Chemogenomics databases.** We have constructed a set of DA-KB and implemented them to the PANDA platform to promote data-sharing and facilitate DA drug target/off target identification and systems pharmacology research for DAR community.

A. Highlights of research accomplishments

- We have Constructed domain-specific knowledge databases for *drug abuse*,[1-2] *hallucinogens*,[3] *Alzheimer's disease*,[4] *cardiovascular disease*[5], and *stem cells*.[6]

B. Representative Research Project

Abstract. To further improve data integration and methods development, we are in development of an integrated computer platform-PANDA (www.cbligand.org/panda) by enriching and centralizing drugs/chemicals data, protein targets, gene expression, metabolism and bioactivity data related to both DA and ND targets (**Figure 1**). The platform's features include cloud computing and sourcing services with integration of our published/established and to-be-developed tools for DA target identification, drug repurposing, and polypharmacology analysis[1, 3-5, 7-9].

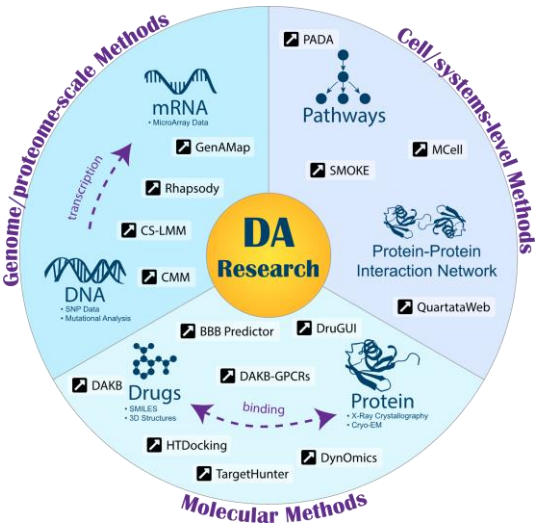


Figure 1. The proposed Platform of Abused-Drugs and Neurological Diseases Association (PANDA) will integrate both established and new chemogenomics tools and databases for drug abuse-related disease research.

Broader Impact of this research on DAR. The in-development PANDA will enhance data sharing, dissemination, and knowledge creation, as well as to boost synergies among the broader DAR community.

2. **GPU-accelerated platform/algorithms/tools for drug target identification for DAR and NDR**

A. Highlights of research accomplishments

- Developed a new platform DAKB-GPCRs:[2] an Integrated computational platform for drug abuse related GPCRs.
- Developed AMBER-based scoring functions for abuse drug-target interactions (ADTI) modeling in DAR[10]
- Applied chemogenomics DBs and computing tools to understand the underlying mechanisms and potential abuse effects of Captagon[11]

B. Representative Research Project 1: DAKB-GPCRs:[2] an Integrated computational platform for drug abuse related GPCRs.

Abstract. Drug abuse (DA) or drug addiction is a complicated brain disorder which is commonly considered as neurobiological impairments caused by both genetic factors and environmental effects. Among DA-related targets, G protein-coupled receptors (GPCRs) play an important role in DA therapy. However, only 39 GPCRs have been published with crystal structures in the recent two decades. In the effort to overcome the limitation of crystal structure and conformational diversity of GPCRs, we built homology models and performed conformational searches by molecular dynamics (MD) simulation. To accelerate and facilitate the drug abuse research, we construct a GPCRs-specific chemogenomics knowledgebase for DA research (DAKB-GPCRs) that implemented with our established and novel chemogenomics tools as well as algorithms for data analyses and visualization. Our established TargetHunter and HTDocking,[2] as well as our novel tools that include target classification and Spider Plot are compiled into the platform as illustrated in **Figure 2**. Our

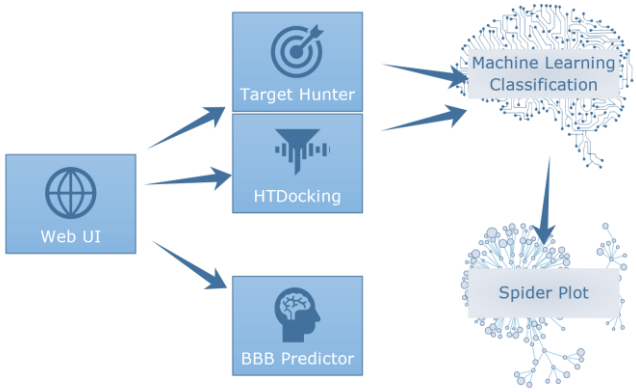


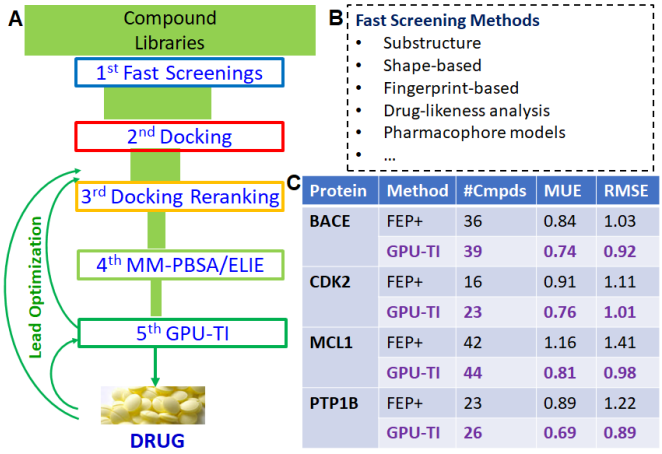
Figure 2. The workflow of DAKB-GPCRs.

DAKB-GPCRs provides the following results for a query compound: (1) blood-brain barrier (BBB) plot via our BBB predictor, (2) docking scores via our HTDocking, (3) similarity score via our TargetHunter,[9] (4) target classification via machine learning methods that utilizes both docking scores and similarity score, and (5) drug-targets interaction network via our Spider Plot**.**

Broader Impact of this research on DAR. To our knowledge, no such domain-specific database is available for the proposed computational applications. Our platform is the first web-based service that integrates DA-related genes, proteins, and drugs for DA research. State-of-the-art computational chemistry/chemoinformatics and machine learning algorithms established in our lab have been implemented for this chemogenomics database, which will help characterize the features of genes, proteins, and drugs in DA study. It will also facilitate new information exchange and data-sharing of knowledge among relevant scientific communities.

C. Representative Research Project 2: Application of physics-based scoring function in binding free energy calculations.

Abstract. By taking advantage of ever-increasing computer power, Core A has developed a successful target-ligand binding free energy calculation depending on both the proper conformational sampling and accuracy of the employed molecular mechanical force field (MMFF) model[12-15]. The MMFF-based free energy calculation is more accurate than the efficient docking scoring functions because of its rigorous theoretic framework which also takes the flexibility of receptors into consideration. MM-PB/GBSA[16-20], linear interaction modeling (**LIE**)[21] and thermodynamic integration (**TI**)[22] are at present widely used MMFF-based methods. **Figure 3** illustrates how those methods will be applied in our strategies of abused drug-target interaction (ADTI) prediction and rational drug design. It is encouraging that AMBER GPU-TI with a general AMBER force field (GAFF)**[23]** has achieved a better performance than FEP+,[24] a commercially available, and costly, software, in reproducing the relative binding free energies for all the four considered protein systems[25-28] as shown in **Figure 3C**. In this renewal, **Core A** will advance, **re-AFFINITY**, a software package which bridges the gap between docking-based and rigorous MMFF-based scoring functions.[10] Moreover, we have successfully expanded the LIE model to **ELIE** (Extended Linear Interaction Energy), by including the polar and nonpolar parts of solvation energies and the entropy term into the parameter search procedures. We plan to develop a set of ELIE models for the **abused drug targets**.[10] re-AFFINITY and ELIE models will be used to re-rank docking poses in the 3rd Step



Protein	Method	#Cmpds	MUE	RMSE
BACE	FEP+	36	0.84	1.03
	GPU-TI	39	0.74	0.92
CDK2	FEP+	16	0.91	1.11
	GPU-TI	23	0.76	1.01
MCL1	FEP+	42	1.16	1.41
	GPU-TI	44	0.81	0.98
PTP1B	FEP+	23	0.89	1.22
	GPU-TI	26	0.69	0.89

Figure 3. Abused drug-target interaction (ADTI) prediction and rational drug strategy with AMBER-based scoring functions. (**A**) Lead identification and optimization strategies; (**B**) fast screening methods, (**C**) Comparison of relative binding free energy calculations (kcal/mol) between FEP+ of Schrodinger and AMBER-GPU TI. MUE and RMSE stands for mean unsigned error and root-mean-square error, respectively.

Broader Impact of this research on DAR. Accurate evaluation of a scoring function helps assess the binding free energy and predict the highest-affinity pose of a ligand binding to its target, which is one of the most important challenges in structure-based drug design for DAR. The developed AMBER FF-based methods have dramatically improved the state-of-the-art in drug discovery targeting drug abuse and neurological disorders.

3. Machine learning (ML) based algorithms and computer models for DAR and NDR

A. Highlights of research accomplishments

- Developed "CSP-Target Mapping",[29] a new GPU-accelerated machine/deep learning-based algorithm tool for target prediction, classification, and mapping, by integration of *TargetHunter*[9] and *HTDocking*[4] algorithms
- Constructed a set of ML-based predictive models using the behavioral and psychological descriptors collected by CEDAR[30-31]
- Predicted orthosteric and allosteric regulations on Cannabinoid receptors using supervised machine learning classifiers[32]

B. Representative Research Project: Computational Systems Pharmacology-Target Mapping for Fentanyl-laced Cocaine Overdose

Abstract. The United States of America is fighting against one of its worst-ever drug crises. Over 900 people a week die from opioid/heroin-related overdoses while millions more suffer from opioid prescription addiction. Recently, drug overdoses caused by the fentanyl-laced cocaine specifically are on the rise. Due to drug synergy and an increase in side effects, polydrug addiction can cause more risk than a single drug. In our recent work, we systematically analyzed the overdose/addiction mechanism of cocaine and fentanyl. Firstly, we applied our established chemogenomics knowledgebase and machine-learning-based methods to map out the potential/known proteins, transporters, metabolic enzymes, and the potential therapeutic target(s) for cocaine and fentanyl. Sequentially, we looked insight into the detail of (1) the addiction to cocaine and fentanyl by binding to the dopamine transporter and the μ opioid receptor (DAT/ μOR); (2) the potential drug-drug interaction of cocaine and fentanyl via p-glycoprotein (P-gp) efflux; (3) the metabolism of cocaine and fentanyl in CYP3A4; (4) the PBPK model for two drugs and their drug-drug interaction in ADME level. Finally, we looked into the detail of JWH133, an agonist of cannabinoid 2-receptor (CB2) with potential therapy for cocaine and fentanyl overdose. All these results provide insight into a better understanding of fentanyl and cocaine polydrug addiction and future drug abuse prevention. **Figure 4** illustrates the complex drug-drug target interactions involved by cocaine and fentanyl.

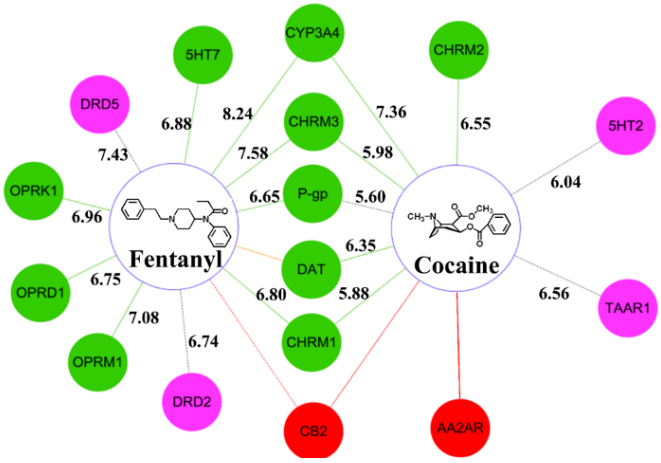


Figure 4. Computational systems pharmacology-target mapping for target proteins, transporters, metabolic enzymes and potential therapeutic targets of cocaine and fentanyl. The green circles and solid lines represented the known targets and interaction for the drugs, while the purple circles and dash lines represented the predicted targets and interaction. In addition, the red circle and solid line represented the known therapeutic target and reported therapeutic effect, while the red dash line represented the predicted therapeutic effect.

Broader Impact of this research on DAR. Fentanyl is one of the most potent opioid substance, has been found to be used as an additive to recreational drugs such as cocaine and heroin to increase their hallucinogen effect in recent years. This phenomenon poses a serious problem on top of the opioid epidemic since the use of only a small amount of fentanyl is required to cause detrimental symptoms such as respiratory depression. In addition, the concordant use of fentanyl and other substance can potentially increase the risk of overdose and other adverse events. Our studies provided detailed information for fentanyl-cocaine polyaddiction from PK and PD aspects. We also suggest CB2 compounds such as JWH133 may have the potential therapeutic effect for both cocaine and fentanyl.

4. Application of molecular systems pharmacology to understand clinical outcomes

A. Highlights of research accomplishments

- Constructed a set of ML-based predictive models using behavioral and psychological descriptors for DA prevention in collaboration with R. Tarter, Director of CEDAR[30-31] (CEDAR center was funded by NIDA)

- Studied the pharmacokinetic and pharmacodynamic drug-drug interactions between opioids and benzodiazepines in order to interpret and understand prescription opioid medication overdose issues[33-34]
- Clinical data mining outcomes analysis and system pharmacology studied synergistic effects of anti-hypertensive drugs and cholinesterase inhibitors on cognitive decline in Alzheimer's patients[35]
- Investigated drug synergy mechanisms of an herbal formula for cardiovascular disease,[5] Conducted systems pharmacology analysis on aspirin/cilostazol to study their therapeutic effects on cardiovascular disease[36] (*cardiotoxicity is one of the issues associated with drug abuse and overdose)

B. Representative Research Project 1: Development of predictive model using behavioral and psychological descriptors for DA prevention

Abstract. We have applied various ML algorithms to analyze DA patient clinical data and also combined systems pharmacology studies to better interpretate outcomes at the molecular level. In collaboration with Dr. Ralph Tarter (Director of the former NIDA-supported Center for Education and Drug Abuse Research (**CEDAR**) at the University of Pittsburgh), we performed data-mining on substance use disorder (**SUD**) data collected by **CEDAR**[37]. The human subjects of this work participated in a longitudinal research study with the funding support from NIH NIDA and conducted by CEDAR on examining the etiology of SUD in families. ***First***, we have successfully defined the clinical outcomes of substance abuse through trajectory analysis of the longitudinal total harm scores, showing that the subjects were classified into either high-risk (**HR**) or low-risk (**LR**) groups (**Figure 5A**). Clearly, there is a significant correlation between the trajectory analysis-based substance abuse risks, and the SUD and DUD clinical diagnoses (**Figure 5B**). From hundreds of etiological variables, 30 were selected by a random forest (RF) algorithm to construct a variety of classification models using different machine learning algorithms (**Figure 5C**). Random forest and Naïve Bayes are the 2 most promising ML methods for building classification models with satisfactory predictive power. For the classifier based on the first-visit data, the area under the curve of receiver operating characteristic (ROC) analysis is 0.71. Although this performance is inferior to those of the classifiers constructed using data of the follow-up visits, it has greater potential since substance abuse prevention measures will be taken when subjects are only at ages of 10 and 11. Two manuscripts were recently accepted for publication in ***Drug and Alcohol Dependence***.[30-31]

Broader Impact of this research on DAR. Those models could make significant contributions to DA prognostication and treatment. For example, we have carried out preliminary studies of a set of classifiers to predict the DA risk group (high or low) to which a subject belongs to, using data obtained in their adolescence (even as young as 10 or 11 years old) through early adulthood with very satisfactory prediction performance.

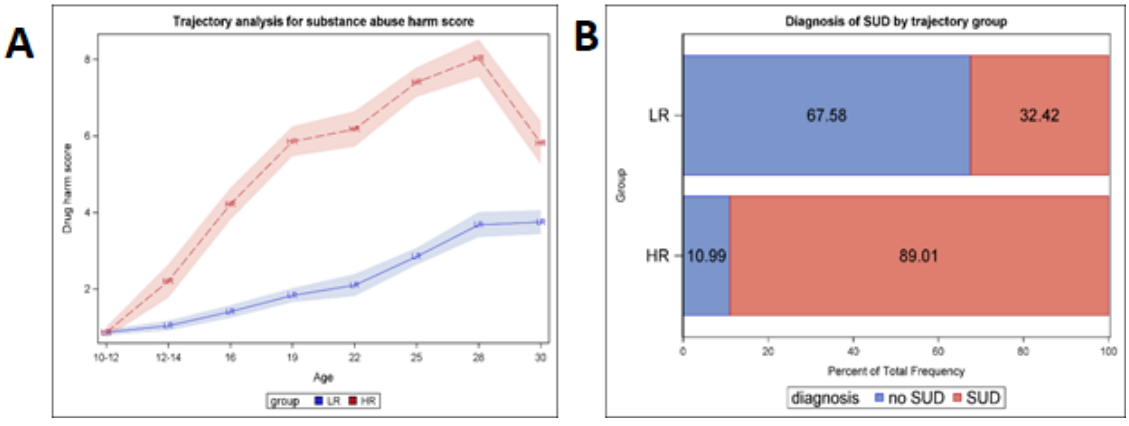


Figure 5. Data-mining of CEDAR-collected human drug abuse (DA) data with multiple ML algorithms. (A) Definition of substance use disorder (SUD) risk through trajectory analysis. Two risk groups are defined: high-risk (**HR**, in red) and low-risk (**LR**, in blue); (B) Summary of correlation between trajectory analysis-based risk classification and clinical diagnoses of substance use disorders.

C. Representative Research Project 2: Drug-drug interaction studies for abused drugs co-administered with benzodiazepines

Abstract. Core A has studied the possible interactions between benzodiazepines (such as diazepam-**DZP**, alprazolam-**APZ**) and opioids (oxycodone-**OXY** and fentanyl- **FEN**) through data mining of the FDA Adverse Event Reporting System (**FAERS**) and other databases, PBPK modeling, and off-target prediction. Pharmacokinetic drug-drug interactions (PK-DDIs) occur when two administered drugs are metabolized by the same enzymes or interact with the same transporters. Our PBPK modeling and simulations[33-34] suggest that OXY and DZP have an interaction only when DZP is overdosed (**Figures 6A, 6B**). On the other hand, the PK-DDI is much more obvious between FEN and APZ (**Figures 6C, 6D).** This result is consistent with a recent report on the antinociceptive tolerance of OXY induced by DZP[38]. To elucidate the mechanism of the DDI between OXY and DZP, we performed CSP-Target studies of DTIs for two drugs using HTDocking. We found that there is an obvious off-target effect for DZP with it acting as an agonist for both μ[39] and κ-opioid[40] receptors, since DZP, OXY, and the known agonists of μ and κ-opioid receptors have comparable docking scores.

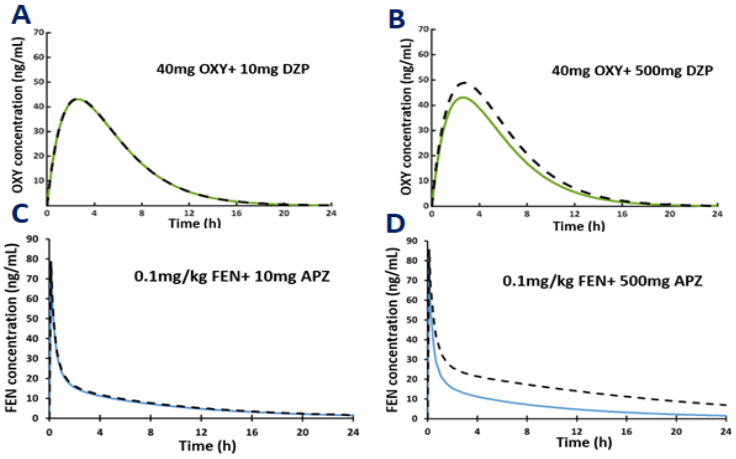


Figure 6. Drug-drug interaction between opioids and benzodiazepines, which is measured by the difference of the AUC of the solid line (opioid only) and dashed-line (co-administrated with a benzodiazepine)**.** (A) 40 mg OXY / 100 mg DZP, (B) 40 mg OXY / 500 mg DZP, (C) 0.1mg/kg FEN / 10 mg APZ, (D) 0.1 mg/kg FEN / 500 mg APZ

Broader Impact of this research on DAR. The PK parameters and covariates identified during population PK/PD analysis will be used to achieve precision medicine of DA treatment for individual patients.

5. Synergy among the Cores and collaboration with FRPs and P/FPs on DAR and NDR

A. Highlights of research accomplishments

- Collaborated with Dr. Bahar (**Core B**) to support Dr. Yong Wan (Northwestern University) to discover novel small molecules for KELF4-PRMT5 with potential therapeutic effects for cancer and neuro-disorders.
- Collaborated with Dr. Wu (**Core C**) for predicting compounds with potential therapy for drug abuse[41]
- Collaborated with former FRP (Y Liu) for discovering novel CB2 inverse agonist for treating kidney fibrosis;[35] and FRP (O Lopez) on the Alzheimer's clinical data analysis[35]
- Supported P/FP PI (I Hernandez) to successfully apply for fellowship and K01 grant (1K01HL142847)
- Supported former P/FP PI (M Torregrossa) to successfully apply for her R01 grant (5R01DA042029)

B. Representative Research Project: A computational strategy for finding novel targets and therapeutic compounds for opioid dependence[41]

Abstract. Opioids are widely used for treating different types of pains, but overuse and abuse of prescription opioids have led to opioid epidemic in the United States. Besides analgesic effects, chronic use of opioid can also cause tolerance, dependence, and even addiction. Effective treatment of opioid addiction remains a big challenge today. Studies on addictive effects of opioids focus on striatum, a main component in the brain responsible for drug dependence and addiction. Some transcription

regulators have been associated with opioid addiction, but relationship between analgesic effects of opioids and dependence behaviors mediated by them at the molecular level has not been thoroughly investigated. In this paper, we developed a new computational strategy that identifies novel targets and potential therapeutic molecular compounds for opioid dependence and addiction in collaboration with **Core C**. We employed several statistical and machine learning techniques and identified differentially expressed genes over time which were associated with dependence-related behaviors after exposure to either morphine or heroin, as well as potential transcription regulators that regulate these genes, using time course gene expression data from mouse striatum. Moreover, our findings revealed that some of these dependence-associated genes and transcription regulators are known to play key roles in opioid-mediated analgesia and tolerance, suggesting that an intricate relationship between opioid-induce pain-related pathways and dependence may develop at an early stage during opioid exposure. Finally, we determined small compounds that can potentially target the dependence-associated genes and transcription regulators. These compounds may facilitate development of effective therapy for opioid dependence and addiction. We also built a database (http://daportals.org) for all opioid-induced dependence-associated genes and transcription regulators that we discovered, as well as the small compounds that target those genes and transcription regulators.

Broader Impact of this research on DAR. In summary, our work here intent to elucidate molecular connections between the analgesic and tolerance-related pain pathways and harmful side effects of opioid use during pain treatment. Despite the general belief that morphine is safe for managing patients with pain, our results suggest that morphine may induce tolerance to analgesia and dependence on the drug in the patients in the very early stage, which may increase the possibility of the same patients to abuse heroin thereafter, since heroin may further induce acute analgesic effects as suggested by our results. Moreover, because heroin can cause both structural and behavioral changes among patients, abusing heroin after morphine may lead to more potent dependence on the drugs among the patients.

OPPORTUNITIES FOR TRAINING AND PROFESSIONAL DEVELOPMENT

Courses Taught by Core A Researchers. The PI Xie and co-investigators participated in teaching the following courses (Total Lecture hours: 50 hrs):

(1) **Drug Discovery, Design & Development Journal Club** (Pharm3048**,** Course Coordinator, 21 hours course time) 2018 Spring.
(2) **Advanced Medicinal Chemistry** (Pharm3032) (3 credits) lectured **8 hours** + 2 hrs practice + 4 hrs practice planned on Computational Medicinal Chemistry, 2018/2019 Fall.
(3) **Foundations in Pharmaceutical Sciences** (Pharm3023) (5 credits) as a course co-coordinator (Part II. Drug Discovery and Design) and lectured **6 hours** + 2 hrs practice + 4 hrs practice planned on computer-aided drug design, 2019 Spring.
(4) **Pharmaceutical Analysis** (PHARM2100, 3 credits), taught 14 hrs lectures + 2 hrs practice + 4 hrs practice planned on NMR application in pharmaceutical sciences, 2019 Spring.
(5) **Computational Systems Pharmacology** (Pharm 3068, 3 credits, a newly developed core course for the PSP program), 2018 fall.
(6) **Pharmacometrics** (Pharm 3069, 3 credit, a newly developed core course for the PSP program), 2019, fall.

Grad students mentoring and research guidance

(1) Dr. Yuanqiang Huang (Visiting Professor from China)
(2) Dr. Jin Cheng (Visiting Professor from China)
(3) Yan Chen (Visiting Professor from China)
(4) Ying Xue (Visiting Professor from China))
(5) Shunqing Xu (Visiting Professor from China)
(6) Weiwei Lin (PharmD student)
(7) Shifan Ma (PhD student)
(8) Ziheng Hu (PhD student)
(9) Changrui Xing (PhD student)

(10) Yankang Jing (PhD student)
(11) Yuemin Bian (PhD student)
(12) Peihao Fan (PhD student)
(13) Jacob Cuyler (PhD student)
(14) Nan Wu (MS student)
(15) Siyi Wang (MS student)
(16) Tianjian Lian (MS student)
(17) Xiguang Qi (MS student)
(18) Yuanyuan Xu (MS student)
(19) Mingzhe Shen (MS student)
(20) Beihong Ji (PhD student)
(21) Shuhan Liu (MS student)
(22) Jingchen Zhai (MS student)
(23) Yuzhao Zhang (MS student)
(24) Dongxiao Hao (Exchange student from China)
(25) Matthew Brock (PharmD student)

Undergraduate students (University of Pittsburgh):

(1) Jack Zhao (Case Western Reserve University, 2018)
(2) Ashna Gupta (Biology)
(3) Emma J Palumbo (Molecular Biology; Genetics)
(4) Erika Thomas (Biological Sciences, Computer Science)
(5) Zechen Wang (Visiting student from Sun Yat-san University, China)

RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST

Publications: During the past year, Core A has published 23 papers in peer-reviewed journals and 13 of them are directly related to DAR (see Products section, below).

Center Programs: Core A has participated in interactive educational enrichment programs supported by the Center. They are briefly described below. Detailed information can be found in the Admin Core Section.

(1) CDAR Annual meeting
(2) Journal Clubs
(3) Joint Lab Meetings
(4) PK/PD modeling software learning clubs
(5) Fellows Training Course/Workshop
(6) CMU/Pitt Educational Opportunities (including Pharmacometrics course)

Scholarly activity (Core A PI Dr. Xiang-Qun Xie)

(1) **Charter Member of Science Board** to the US FDA Advisory Committee, the United States, the Food and Drug Administration (FDA) (2014-2018)
(2) **Editorial Advisory Board Member** of American Association of Pharmaceutical Scientists (AAPS) Journal 2015-present)

Invited Talks (Core A PI Xiang-Qun Xie and Other Co-Is)

(1) (Invited speaker) Xie, XQ, 2018 Nov 19 “Pharmacometrics & System Pharmacology (PSP) Program for Drug Discovery”, Macau University of Science and Technology (MUST;澳門科技大學), China.
(2) (Invited speaker) **Xie, XQ**, 2018 Nov 1-3 “GPCRs-KB. An Integrated Platform of GPCRs Computational Chemogenomics KnowledgeBase for System Pharmacology Drug Discovery” in

2018 International Conference of Frontiers in Precision & Translational Medicine (PTMC2018)", Peking University Health Science Center, Beijing, China.

(3) (Invited speaker) **Xie, XQ**, 2018 Nov 12 "GPCRs-Drug Abuse. An Integrated Platform of GPCRs Computational Chemogenomics KnowledgeBase for System Pharmacology Drug Discovery" SYSU, Guangzhou, China

(4) (Invited speaker) Yuemin Bian **XiangQun Xie** "Chemogenomics-knowledge based systems pharmacology analysis and integrated in silico simulation on cannabidiol (CBD)" Chemistry & Pharmacology of Drug Abuse, Boston, United States, 2018.

(5) (Invited speaker) Junmei Wang, "Update on GAFF/Antechamber Development" AMBER Developers' Meeting, Tampa, FL, 2019 March

(6) (Invited speaker) Xibing He, "GAFF3 – the Next Generation of AMBER All-Atom Force Fields" AMBER Developers' Meeting, Tampa, FL, 2019 March

(7) (Invited speaker) Viet Man, "Nonequilibrium MD simulations with laser and bubbles", AMBER Developers' Meeting, Tampa, FL, 2019 March

(8) Junmei Wang, "Rational design of small molecules inhibiting Amyloid Beta aggregation" ACS Meeting, Orlando FL, March 31- April 4, 2019

(9) (Invited speaker) Junmei Wang, International symposium of "Free energy calculations: entering the fourth decade of adventure in chemistry and biophysics", Santa Fe NM, June 16-21, 2019.

(10) Peihao Fan, Nanyi Wang, Lirong Wang, **Xiang-Qun Xie** "Computational Approach for Autophagy and Apoptosis Specific Knowledgebases and Chemogenomics Systems Pharmacology Drug Research" in 27th AAPS Conference.

Conference Posters (Xie's Core A):

(1) **S. Ma**, L. Wang, J. J. Jun, X-Q. Xie.* A Comprehensive Gene Signature Analysis for Multiple Myeloma. *UPITT Science Conference 2019*, Pittsburgh, PA, United States.

(2) **Yuemin Bian**, X-Q Xie "The effects and mechanism of α-mangostinderivatives on the Alzheimer's disease model of rat: a combination of experimental study and computational systems pharmacology analysis" *UPITT Science Conference 2018*, Pittsburgh, United States, 2018.

(3) **Ji, B**. "Mechanism of drug-drug interaction between Opioids and Benzodiazepines", *Pitt-Pharmacy Graduate Program Poster Presentation and Awards Dinner*, Oct., 9, 2018

(4) **Ji, B**. "Mechanism of drug-drug interaction between Opioids and Benzodiazepines", *UPITT Science Conference 2018*, Oct. 18, 2018

(5) **Fan, P**. "Computational Approach for autophagy and apoptosis specific knowledgebases and chemogenomics systems pharmacology drug research", *UPITT Science Conference 2018*, Oct. 18, 2018

(6) **Shifan Ma**, LiRong Wang, Xiaang-Qun Xie. "A meta-analysis comparing multiple myeloma gene signatures and gene or chemical perturbation induced gene profiles" *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(7) **Changrui Xing**, Youwen Zhuang, Cheng Zhang, Eric Xu, Xiang-Qun Xie. "Cannabinoid Receptor CB2 Structure and CB2/Gi Signaling Mechanisms." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(8) **Peihao Fan**, Nanyi Wang, LiRong Wang, Xiang-Qun Xie. "Computational Approach for Autophagy And Apoptosis Specific Knowledgebases and Chemogenomics Systems Pharmacology Drug Research." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(9) Yuemin Bian, **Xibing He**, Lirong Wang, Junmei Wang, Xiang-Qun Xie. "Computational systems pharmacology analysis of cannabidiol: a combination of chemogenomics knowledgebase network analysis and integrated in silico modeling and simulation." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(10) **Siyi Wang,** Jin Cheng, Weiwei Lin, Nan Wu, Yuanqiang Wang, Maozi Chen, Xiang-Qun (Sean) Xie, and Zhiwei Feng. "Computational System Pharmacology-Target Mapping (CSP- Target Mapping) For Fentanyl-laced Cocaine Overdose." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(11) **Viet H. Man.** "Promising applications of ultrasound and infrared laser on brain therapies: a theoretical study." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(12) **Shuhan Liu,** Xibing He, Viet Hoang Man, Beihong Ji, Junmei Wang. "New Application of *In Silico* Methods in Identifying Key Components of Anti-Cancer Herbal Formulation YIV-906." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(13) Beihong Ji, Ying Xue, Yuanyuan Xu, Zhaojia Zhang, Shuhan Liu, Albert H Gough, Junmei Wang*, Lirong Wang*, Xiang-Qun Xie. (Presented by **Mingzhe Shen**) "Physiologically-based Pharmacokinetics Modeling of Drug-Drug Interaction between Oxycodone and Diazepam." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

(14) **Xibing He**, Junmei Wang. "Protocols of Calculating Protein-Ligand Binding Free Energies Using AMBER Programs and Force Fields." *UPitt SOP Graduate Student & Post Doc Retreat and Awards Dinner*, Seven Springs, PA, June 20, 2019

Support of New Funding Efforts of PIs for FRPs (Core A PI Xiang-Qun Xie)

The following grant applications were provided joint efforts or collaboration efforts by Core A and team.

Funded

- **P30 DA035778 (Director: Xie)** 09/01/2014-08/31/2019 (**1 Year No Cost Extension**) 09/01/2020-08/31/2025 **(Renewal Pending)**
  NIDA Center of Excellence of Computational Drug Abuse Research (CDAR), NIH.
  The overall goal of the Computational Drug Abuse Research (CDAR) Center is to advance state of-the art computational technologies for research toward the prevention and treatment of drug abuse (DA) and DA-related diseases.

- **DOD W81XWH-16-1-049 (Xie)** 09/01/2016-08/30/19 (**1 Year No Cost Extension**)
  Chemogenomics Systems Pharmacology Approach for TBI and AD Research
  Alzheimer's disease (AD) is the most common form of dementia. It is a condition associated with memory loss (particularly episodic memory), a slow decline in cognitive ability, and behavioral and physical disability, ultimately resulting in death. Military personnel and other individuals who suffer from traumatic brain injury (TBI) face an increased risk for developing several long-term health problems including AD-like dementia, aggression, memory loss, depression, and symptoms similar to those of AD. We have already obtained significant results through application of computational approaches and have published these findings (*J Neurotrauma* 2018 Sep 6. doi: 10.1089/neu.2018.5757). We are now preparing to perform animal studies and clinical data analysis to find the mechanisms that cause TBI-induced AD. *This funded proposal represents a strong collaboration between Core A (Xie) and* ***Dr. Oscar Lopez*** *and* ***Dr. Dandan Sun.***

- **NIH R01 (Xie and L Wang)**
  Personalized Medicine for Alzheimer's Disease with Common Comorbidities.
  PI: Dr. Xiang-Qun Xie
  Co-PI: Dr Lirong Wang, Dr. Oscar Lopez
  Co-I: Dr. Gregory Cooper
  Initially Submitted: June 5, 2018

- **3R01MH113857 - 02W1** (Price, Rebecca) 12/01/18-06/30/22 0.7 calendar
  NIH/NIMH $46,648
  Improving Precision of Ketamine Metabolite Assays
  Role: Pharmacokinetics Expert (**Junmei Wang**)
  This project seeks to identify the neural and cognitive changes that accompany rapid relief from depressive symptoms following intravenous ketamine.

- **QUMP** ***I***5UL1TR001857 (**Junmei Wang**) 02/01/2019 – 01/31/2020
  PITT CTSI/NIH $10,000
  Quantitatively predict drug-drug interactions between oxycodone and other drugs
  Role: Principal Investigator
- **NIH 1R01MH116046-01A1** (PIs: Sweet, Kofler and **Wang L**) 09/25/2018-06/30/2023
  Synaptic Resilience to Psychosis in Alzheimer Disease

Pending

- **Thome Memorial Foundation Awards Program in Alzheimer's Disease Drug Discovery Research (Xie)** 02/1/2020-01/31/2022
  "Novel SQSTM-1/p62 modulators for the treatment of Alzheimer's Disease"
  Funds Requested: $500,000
  Co-Is: Terence McGuire, Jaden Jun, and Zhiwei Feng
  Full Proposal Submitted: September 9, 2019

- **Proposal #7910508 (Wang, JM)** 04/01/2019-03/31/2022
  National Science Foundation $947,245
  "Molecular Mechanics force field toolkit for studying protein-ligand interactions"

- **NIH R01 (Xie)**
  "CB2 Structure and CB2/Gi signaling mechanisms: Insight into New CB2 Drug Discovery"
  Date Submitted: October 5, 2019

- **NIH R01 (Xie)**
  "INK4C-Targeted Small-Molecule Inhibitors"
  Co-Is: Dr. Jaden Jun, Dr. Terence McGuire, Dr. Zhiwei Feng
  Date Submitted: October 5, 2019

- ***NASA Space CASIS (The Center for the Advancement of Science in Space) (Xie)***
  "Crystalization of CB2"
  Funds Requested: $70,000
  Submitted: May, 2018
  Status: ***Funding appears to have been approved*** and we are actively engaged in planning the microgravity experiment for the rocket launch in spring of 2020.

PLAN FOR NEXT PERIOD

We will continue to improve our existing programs and develop new computational chemogenomics databases and tools (**Aims 1 and 2**), and to develop computational models to predict DA clinical outcomes (**Aim 3**). The new DBs, computational tools and models will be applied to the research of our selected FRPs and pilot projects that are on-going collaborations. We will focus on (but not limited to) the defined plans below:

1. **To construct a new Platform for Abused-Drugs and Neurological Diseases Association (PANDA).** We will enhance data sharing, dissemination, and KB creation as well as boost synergies among the broader DAR community by integrating the chemogenomics KBs of abused drugs with the metabolism pathway/genetic information of DA-associated diseases. We will build the integrated PANDA system to centralize all DA data and tools from all research Cores. Core A will make data accessible through PANDA on: (1) Chemogenomics of DA and **ND**s-related targets and compounds jointly with Computational and Systems Biology Core for DA (**Core B**);[1, 5, 9] (2) gene expression induced by drugs of abuse in cooperation with the Computational Genomics for DA (**Core C**); and (3) metabolism information on **drug-drug interactions** (**DDI**) related to DA. We will improve our existing webtools (www.cbligand.org/cdar) by implementing a user-friendly interface and GPU high-performance computations.

2. **To advance our computing algorithms/tools for modeling abused drugs and DA targets interactions (DTI) and to predict DA treatments against polyaddiction by deep/machine learning (DL/ML) methods.** We will focus on: (1) advancing our novel **c**hemogenomics **s**ystems **p**harmacology (CSP)-Target Mapping program, a new GPU-accelerated ML/DL computing algorithm[8,

[30-32, 42-44] to map out DA targets related to cell signaling by integration of TargetHunter[61] and HTDocking[34] in synergy with **Core C**; (2) modeling the energetics of DA drug-target interactions using the AMBER force field-based scoring functions (AMBER-FFSF)[19, 23, 45] to be used in conjunction with the DruGUI software[46] (for druggability simulations) developed by **Core B**; (3) developing a novel scoring-function methodology for characterizing DA drug/targets complexes, called molecular complex characterizing system (MCCS); and (4) applying these tools to predict DTIs and to discover novel small-molecules for DA/ND targets.[1, 47-49]

3. **To advance our developed ML-based models to predict DA clinical outcomes in combination with systems pharmacology method for assisting in DA prevention.** To achieve this goal, we will focus on: (1) analyzing the clinical data of juveniles, for example, those collected by the Center for Education and Drug Abuse Research (CEDAR), using our established cutting-edge ML/DL computing tools;[7-8, 44, 49-51] (2) applying both the clinical data-mining and computational systems pharmacology tools to model and predict DDIs and DTIs for understanding the molecular basis of DA-related clinical outcomes; and (3) utilizing physiology-based pharmacokinetics (PBPK) modeling and pharmacometrics approaches to quantitatively study DDIs and the underlying mechanisms from the perspectives of drug metabolism and drug synergism.

PRODUCTS

Publications:

1. Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X. Q., DAKB-GPCRs: An Integrated Computational Platform for Drug Abuse Related GPCRs. *Journal of chemical information and modeling* **2019,** *59* (4), 1283-1289.
2. Jing, Y.; Hu, Z.; Fan, P.; Xue, Y.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder. *Drug And Alcohol Depend* **2019**, In Press.
3. Hu, Z.; Jing, Y.; Xue, Y.; Fan, P.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X. Q., Analysis of substance use and its outcomes by machine learning: II. Derivation and prediction of the trajectory of substance use severity. *Drug and Alcohol Dependence* **2019**, In press.
4. Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X. Q.; Feng, Z., Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose. *ACS Chem Neurosci* **2019,** *10* (8), 3486-3499.
5. Wu, X.; Xie, S.; Wang, L.; Fan, P.; Ge, S.; Xie, X. Q.; Wu, W., A computational strategy for finding novel targets and therapeutic compounds for opioid dependence. *PLoS One* **2018,** *13* (11), e0207027.
6. Zhou, L.; Zhou, S.; Yang, P.; Tian, Y.; Feng, Z.; Xie, X. Q.; Liu, Y., Targeted inhibition of the type 2 cannabinoid receptor is a novel approach to reduce renal fibrosis. *Kidney Int* **2018,** *94* (4), 756-772.
7. Bian, Y.; Jing, Y.; Wang, L.; Ma, S.; Jun, J. J.; Xie, X. Q., Prediction of Orthosteric and Allosteric Regulations on Cannabinoid Receptors Using Supervised Machine Learning Classifiers. *Mol Pharm* **2019,** *16* (6), 2605-2615.
8. Bian, Y. M.; He, X. B.; Jing, Y. K.; Wang, L. R.; Wang, J. M.; Xie, X. Q., Computational systems pharmacology analysis of cannabidiol: a combination of chemogenomics-knowledgebase network analysis and integrated in silico modeling and simulation. *Acta Pharmacol Sin* **2019,** *40* (3), 374-386.
9. Ge, H.; Bian, Y.; He, X.; Xie, X. Q.; Wang, J., Significantly different effects of tetrahydroberberrubine enantiomers on dopamine D1/D2 receptors revealed by experimental study and integrated in silico simulation. *J Comput Aided Mol Des* **2019,** *33* (4), 447-459.
10. Man, V. H.; He, X.; Derreumaux, P.; Ji, B.; Xie, X. Q.; Nguyen, P. H.; Wang, J., Effects of All-Atom Molecular Mechanics Force Fields on Amyloid Peptide Assembly: The Case of Abeta16-22 Dimer. *J Chem Theory Comput* **2019,** *15* (2), 1440-1452.
11. Wang, Y. Q.; Lin, W. W.; Wu, N.; Wang, S. Y.; Chen, M. Z.; Lin, Z. H.; Xie, X. Q.; Feng, Z. W., Structural insight into the serotonin (5-HT) receptor family by molecular docking, molecular dynamics simulation and systems pharmacology analysis. *Acta Pharmacol Sin* **2019,** *40*, 1138–1156.

12. Wu, N.; Feng, Z.; He, X.; Kwon, W.; Wang, J.; Xie, X. Q., Insight of Captagon Abuse by Chemogenomics Knowledgebase-guided Systems Pharmacology Target Mapping Analyses. *Sci Rep* **2019,** *9* (1), 2268.
13. Ji, B.; Liu, S.; Xue, Y.; He, X.; Man, V. H.; Xie, X. Q.; Wang, J., Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation. *Drugs R & D* **2019,** *19*.
14. Fan, P.; Wang, N.; Wang, L.; Xie, X. Q., Autophagy And Apoptosis Specific Knowledgebases-Guided Systems Pharmacology Drug Research. *Curr Cancer Drug Targets* **2019**, E-pub Ahead of Print.
15. Ma, S.; Attarwala, I. Y.; Xie, X. Q., SQSTM1/p62: A Potential Target for Neurodegenerative Disease. *ACS Chem Neurosci* **2019,** *10* (5), 2094-2114.
16. Wang, L.; Ma, S.; Hu, Z.; McGuire, T. F.; Xie, X. S., Chemogenomics Systems Pharmacology Mapping of Potential Drug Targets for Treatment of Traumatic Brain Injury. *J Neurotrauma* **2019,** *36* (4), 565-575.
17. Hu, Z.; Wang, L.; Ma, S.; Kirisci, L.; Feng, Z.; Xue, Y.; Klunk, W. E.; Kamboh, M. I.; Sweet, R. A.; Becker, J.; Lv, Q.; Lopez, O. L.; Xie, X. Q., Synergism of antihypertensives and cholinesterase inhibitors in Alzheimer's disease. *Alzheimers Dement (N Y)* **2018,** *4*, 542-555.
18. He, X. M., V. H.; Ji, B.; Xie, X.Q.; Wang, J. , Calculate protein-ligand binding affinities with the extended linear interaction energy method: Application on the cathepsin S set in the D3R grand challenge 3. *J Comput Aided Mol Des* **2019,** (33), 105-117.
19. Man, V. H.; Li, M. S.; Wang, J.; Derreumaux, P.; Nguyen, P. H., Nonequilibrium atomistic molecular dynamics simulation of tubular nanomotor propelled by bubble propulsion. *J Chem Phys* **2019,** *151* (2), 024103.
20. Man, V. H.; Truong, P. M.; Li, M. S.; Wang, J.; Van-Oanh, N. T.; Derreumaux, P.; Nguyen, P. H., Molecular Mechanism of the Cell Membrane Pore Formation Induced by Bubble Stable Cavitation. *J Phys Chem B* **2019,** *123* (1), 71-78.
21. Wang, J.; Ge, Y.; Xie, X. Q., Development and Testing of Druglike Screening Libraries. *Journal of chemical information and modeling* **2019,** *59* (1), 53-65.
22. Zhou, Z.; Feng, Z.; Hu, D.; Yang, P.; Gur, M.; Bahar, I.; Cristofanilli, M.; Gradishar, W. J.; Xie, X. Q.; Wan, Y., A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine* **2019,** *44*, 98-111.
23. Man, V. H.; Li, M. S.; Wang, J. M.; Derreumaux, P.; Nguyen, P. H., Interaction mechanism between the focused ultrasound and lipid membrane at the molecular level. *J Chem Phys* **2019,** *150* (21).

Computation Algorithms. Models and Tools:

(1) **TargetHunter**: Ligand-based target prediction tools with BioassayGeoMap function integrated. http://www.cbligand.org/TargetHunter
(2) **HTDocking**: Structure-based target prediction tools for drug repurpose research. http://www.cbligand.org/HTDocking/
(3) **BBB Predictor**: Machine learning algorithms- AdaBoost and SVM-based tool is designed for predicting the permeability of blood-brain barrier (BBB) for compounds. http://www.cbligand.org/BBB/
(4) **GPU-Accelerated Compound Library Comparison**: Modern graphics process units (GPU) – based parallel computing, millions of compound comparison can be accomplished in a few seconds. http://www.cbligand.org/gpu/
(5) **LiCABEDS:** Ligand Classifier of Adaptively Boosting Ensemble Decision Stumps (LiCABEDS), based on machine learning algorithm for compound classification prediction (www.CBLigand.org/LiCABEDS)
(6) **PAINS-Remover:** A Computer-Aided Drug Design (CADD) tool to identify the false positives in experimental high-throughput screening data. (www.CBLigand.org/FPR)
(7) **Mol-Prop:** GPU-accelerated molecular property calculation. (www.CBLigand.org/gpu)
(8) **MMFFT**: A user-friendly web toolkit for generating molecular mechanical force field (MMFF) models for arbitrary chemicals. (https://mulan.pharmacy.pitt.edu/mmfft)
(9) **re-Affinity**: A software tool to re-rank docking poses using the MM-PB/GBSA scoring functions (https://mulan.pharmacy.pitt.edu)

Chemogenomics Databases:

(1) **CBID:** Cannabinoid molecular information database http://www.cbligand.org/cbid
(2) **AIzPlatform**: Chemogenomics Database for Alzheimer's Disease, (http://www.cbligand.org/AD)
(3) **Hallucinogen:** Chemogenomics Database for Hallucinogen Research (http://www.cbligand.org/hallucinagen)
(4) **StemCellCKB**: Chemogenomics Database for Stem Cell Research (http://cbligand.org/StemCellCKB)
(5) **CVDPlatform**: Chemogenomics Database for Cardiovascular Disease (http://www.cbligand.org/CVD)
(6) **DAKB**: Chemogenomics Database for Drug abuse Research is designed for facilitating data-sharing and information exchange among scientific research communities for drug abuse, including genes, proteins, small molecules and signal pathways, with online structure search functions and data analysis tools implemented. http://www.cbligand.org/CDAR
(7) **Hallucinogen**: Chemogenomics KB for Hallucinogen research (www.CBLigand.org/hallucinogen)
(8) **CVD**: Chemogenomics KB for cardiovascular diseases (www.cbligand.org/CVD)
(9) **TBI**: Chemogenomics KB for traumatic brain injury (www.CBLigand.org/TBI)

More technologies for core A can be found in our CCGS center. http://www.cbligand.org/CCGS/

REFERENCES CITED

1. Xie, X. Q.; Wang, L.; Liu, H.; Ouyang, Q.; Fang, C.; Su, W., Chemogenomics knowledgebased polypharmacology analyses of drug abuse related G-protein coupled receptors and their ligands. *Front Pharmacol* **2014,** *5*, 3.
2. Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X. Q., DAKB-GPCRs: An Integrated Computational Platform for Drug Abuse Related GPCRs. *J Chem Inf Model* **2019,** *59* (4), 1283-1289.
3. Xu, X.; Ma, S.; Feng, Z.; Hu, G.; Wang, L.; Xie, X. Q., Chemogenomics knowledgebase and systems pharmacology for hallucinogen target identification-Salvinorin A as a case study. *J Mol Graph Model* **2016,** *70*, 284-295.
4. Liu, H.; Wang, L.; Lv, M.; Pei, R.; Li, P.; Pei, Z.; Wang, Y.; Su, W.; Xie, X. Q., AlzPlatform: an Alzheimer's disease domain-specific chemogenomics knowledgebase for polypharmacology and target identification research. *J Chem Inf Model* **2014,** *54* (4), 1050-1060.
5. Zhang, H.; Ma, S.; Feng, Z.; Wang, D.; Li, C.; Cao, Y.; Chen, X.; Liu, A.; Zhu, Z.; Zhang, J.; Zhang, G.; Chai, Y.; Wang, L.; Xie, X.-Q., Cardiovascular disease chemogenomics knowledgebase-guided target identification and drug synergy mechanism study of an herbal formula. *Sci Rep* **2016,** *6*, 33963.
6. Zhang, Y.; Wang, L.; Feng, Z.; Cheng, H.; McGuire, T. F.; Ding, Y.; Cheng, T.; Gao, Y.; Xie, X. Q., StemCellCKB: An Integrated Stem Cell-Specific Chemogenomics KnowledgeBase for Target Identification and Systems-Pharmacology Research. *J Chem Inf Model* **2016,** *56* (10), 1995-2004.
7. Myint, K.; Wang, L.; Tong, Q.; Xie*, X.-Q., Fingerprint-based artificial neural networks qsar (fann-qsar) for ligand biological activity predictions. *Mol Pharm* **2012,** *1260*, 149-164.
8. Wang, L.; Ma, C.; Wipf, P.; Xie, X. Q., Linear and Nonlinear Support Vector Machine for the Classification of Human 5-Ht1a Ligand Functionality. *Mol Inform* **2012,** *31* (1), 85-95.
9. Wang, L.; Ma, C.; Wipf, P.; Liu, H.; Su, W.; Xie, X.-Q., Targethunter: An in Silico Target Identification Tool for Predicting Therapeutic Potential of Small Organic Molecules Based on Chemogenomic Database. *AAPS J* **2013,** *15* (2), 395-406.
10. He, X. M., V. H.; Ji, B.; Xie, X.Q.; Wang, J. , Calculate protein-ligand binding affinities with the extended linear interaction energy method: Application on the cathepsin S set in the D3R grand challenge 3. *J Comput Aided Mol Des* **2019,** (33), 105-117.

11. Wu, N.; Feng, Z.; He, X.; Kwon, W.; Wang, J.; Xie, X. Q., Insight of Captagon Abuse by Chemogenomics Knowledgebase-guided Systems Pharmacology Target Mapping Analyses. *Sci Rep* **2019,** *9* (1), 2268.
12. Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* **2013,** *9* (9), 3878-3888.
13. Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., NAMD2: Greater scalability for parallel molecular dynamics. *J Comput Phys* **1999,** *151* (1), 283-312.
14. Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput* **2012,** *8* (5), 1542-1555.
15. Salomon-Ferrer, R.; Case, D. A.; Walker, R. C., An overview of the Amber biomolecular simulation package. *Wires Comput Mol Sci* **2013,** *3* (2), 198-210.
16. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* **2011,** *51* (1), 69-82.
17. Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M., Validation and use of the MM-PBSA approach for drug discovery. *J Med Chem* **2005,** *48* (12), 4040-4048.
18. Swanson, J. M.; Henchman, R. H.; McCammon, J. A., Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys J* **2004,** *86* (1 Pt 1), 67-74.
19. Wang, J. M.; Hou, T. J.; Xu, X. J., Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Curr Comput-Aid Drug* **2006,** *2* (3), 287-306.
20. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T., End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem Rev* **2019,** *119* (16), 9478-9508.
21. Su, Y.; Gallicchio, E.; Das, K.; Arnold, E.; Levy, R. M., Linear Interaction Energy (LIE) Models for Ligand Binding in Implicit Solvent: Theory and Application to the Binding of NNRTIs to HIV-1 Reverse Transcriptase. *J Chem Theory Comput* **2007,** *3* (1), 256-277.
22. Lee, T. S.; Hu, Y.; Sherborne, B.; Guo, Z.; York, D. M., Toward Fast and Accurate Binding Affinity Prediction with pmemdGTI: An Efficient Implementation of GPU-Accelerated Thermodynamic Integration. *J Chem Theory Comput* **2017,** *13* (7), 3077-3084.
23. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of computational chemistry* **2004,** *25* (9), 1157-1174.
24. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R., Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J Am Chem Soc* **2015,** *137* (7), 2695-2703.
25. Cumming, J. N.; Smith, E. M.; Wang, L.; Misiaszek, J.; Durkin, J.; Pan, J.; Iserloh, U.; Wu, Y.; Zhu, Z.; Strickland, C.; Voigt, J.; Chen, X.; Kennedy, M. E.; Kuvelkar, R.; Hyde, L. A.; Cox, K.; Favreau, L.; Czarniecki, M. F.; Greenlee, W. J.; McKittrick, B. A.; Parker, E. M.; Stamford, A. W., Structure based design of iminohydantoin BACE1 inhibitors: identification of an orally available, centrally active BACE1 inhibitor. *Bioorg Med Chem Lett* **2012,** *22* (7), 2444-2449.
26. Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J., N2-substituted O6-cyclohexylmethylguanine derivatives: potent inhibitors of cyclin-dependent kinases 1 and 2. *J Med Chem* **2004,** *47* (15), 3710-3722.
27. Wilson, D. P.; Wan, Z. K.; Xu, W. X.; Kirincich, S. J.; Follows, B. C.; Joseph-McCarthy, D.; Foreman, K.; Moretto, A.; Wu, J.; Zhu, M.; Binnun, E.; Zhang, Y. L.; Tam, M.; Erbe, D. V.; Tobin, J.; Xu, X.; Leung, L.; Shilling, A.; Tam, S. Y.; Mansour, T. S.; Lee, J., Structure-based optimization of protein tyrosine phosphatase 1B inhibitors: From the active site to the second phosphotyrosine binding site. *Journal of Medicinal Chemistry* **2007,** *50* (19), 4681-4698.

28. Friberg, A.; Vigil, D.; Zhao, B.; Daniels, R. N.; Burke, J. P.; Garcia-Barrantes, P. M.; Camper, D.; Chauder, B. A.; Lee, T.; Olejniczak, E. T.; Fesik, S. W., Discovery of Potent Myeloid Cell Leukemia 1 (Mcl-1) Inhibitors Using Fragment-Based Methods and Structure-Based Design. *J Med Chem* **2013,** *56* (1), 15-30.
29. Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X. Q.; Feng, Z., Computational Systems Pharmacology-Target Mapping for Fentanyl-Laced Cocaine Overdose. *ACS Chem Neurosci* **2019,** *10* (8), 3486-3499.
30. Jing, Y.; Hu, Z.; Fan, P.; Xue, Y.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X.-Q., Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder. *Drug And Alcohol Depend* **2019**, In Press.
31. Hu, Z.; Jing, Y.; Xue, Y.; Fan, P.; Wang, L.; Tarter, R.; Kirisci, L.; Vanyukov, M. M.; Wang, J.; Xie, X. Q., Analysis of substance use and its outcomes by machine learning: II. Derivation and prediction of the trajectory of substance use severity. *Drug and Alcohol Dependence* **2019**, In press.
32. Bian, Y.; Jing, Y.; Wang, L.; Ma, S.; Jun, J. J.; Xie, X. Q., Prediction of Orthosteric and Allosteric Regulations on Cannabinoid Receptors Using Supervised Machine Learning Classifiers. *Mol Pharm* **2019,** *16* (6), 2605-2615.
33. Ji, B.; Xue, Y.; Xu, Y.; Zhang, Z.; Gough, A. H.; Wang, J.; Wang, L.; Xie, X.-Q., Pharmacokinetics Modeling of Drug-Drug Interaction Between Oxycodone and Diazepam *Eur J Pharmceutical Sciences*, Submitted.
34. Ji, B.; Liu, S.; Xue, Y.; He, X.; Man, V. H.; Xie, X. Q.; Wang, J., Prediction of drug-drug interactions between opioids and overdosed benzodiazepines using physiologically-based pharmacokinetic (PBPK) modeling and simulation. *Drugs R & D* **2019**, In press.
35. Hu, Z.; Wang, L.; Ma, S.; Kirisci, L.; Feng, Z.; Xue, Y.; Klunk, W. E.; Kamboh, M. I.; Sweet, R. A.; Becker, J.; Lv, Q.; Lopez, O. L.; Xie, X. Q., Synergism of antihypertensives and cholinesterase inhibitors in Alzheimer's disease. *Alzheimers Dement (N Y)* **2018,** *4*, 542-555.
36. Xue, Y.; Feng, Z. W.; Li, X. Y.; Hu, Z. H.; Xu, Q.; Wang, Z.; Cheng, J. H.; Shi, H. T.; Wang, Q. B.; Wu, H. Y.; Xie, X. Q.; Lv, Q. Z., The efficacy and safety of cilostazol as an alternative to aspirin in Chinese patients with aspirin intolerance after coronary stent implantation: a combined clinical study and computational system pharmacology analysis. *Acta Pharmacol Sin* **2018,** *39* (2), 205-212.
37. Vanyukov, M. M.; Kirisci, L.; Moss, L.; Tarter, R. E.; Reynolds, M. D.; Maher, B. S.; Kirillova, G. P.; Ridenour, T.; Clark, D. B., Measurement of the risk for substance use disorders: phenotypic and genetic analysis of an index of common liability. *Behav Genet* **2009,** *39* (3), 233-244.
38. Gonek, M.; Akbarali, H. I.; Henderson, G.; Dewey, W. L., Reversal of oxycodone and hydrocodone tolerance by diazepam. *Brain Res* **2017,** *1674*, 84-90.
39. Huang, W. J.; Manglik, A.; Venkatakrishnan, A. J.; Laeremans, T.; Feinberg, E. N.; Sanborn, A. L.; Kato, H. E.; Livingston, K. E.; Thorsen, T. S.; Kling, R. C.; Granier, S.; Gmeiner, P.; Husbands, S. M.; Traynor, J. R.; Weis, W. I.; Steyaert, J.; Dror, R. O.; Kobilka, B. K., Structural insights into mu-opioid receptor activation. *Nature* **2015,** *524* (7565), 315-321.
40. Che, T.; Majumdar, S.; Zaidi, S. A.; Ondachi, P.; McCorvy, J. D.; Wang, S.; Mosier, P. D.; Uprety, R.; Vardy, E.; Krumm, B. E.; Han, G. W.; Lee, M. Y.; Pardon, E.; Steyaert, J.; Huang, X. P.; Strachan, R. T.; Tribo, A. R.; Pasternak, G. W.; Carroll, F. I.; Stevens, R. C.; Cherezov, V.; Katritch, V.; Wacker, D.; Roth, B. L., Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell* **2018,** *172* (1-2), 55-67.
41. Wu, X.; Xie, S.; Wang, L.; Fan, P.; Ge, S.; Xie, X. Q.; Wu, W., A computational strategy for finding novel targets and therapeutic compounds for opioid dependence. *PLoS One* **2018,** *13* (11), e0207027.
42. Ma, C.; Wang, L.; Xie*, X.-Q., GPU Accelerated Chemical Similarity Calculation for Compound Library Comparison. *Journal of Chemical Information and Modeling* **2011,** *51* (7), 1521-1527.
43. Myint, K. Z.; Xie, X. Q., Ligand biological activity predictions using fingerprint-based artificial neural networks (FANN-QSAR). *Methods Mol Biol* **2015,** *1260*, 149-164.
44. Jing, Y.; Bian, Y.; Hu, Z.; Wang, L.; Xie, X. Q., Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era. *AAPS J* **2018,** *20* (3), 58.
45. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **2006,** *25* (2), 247-260.
46. Bakan, A.; Nevins, N.; Lakdawala, A. S.; Bahar, I., Druggability Assessment of Allosteric Proteins by Dynamics Simulations in the Presence of Probe Molecules. *J Chem Theory Comput* **2012,** *8* (7), 2435-2447.

47. Feng, Z.; Alqarni, M. H.; Yang, P.; Tong, Q.; Chowdhury, A.; Wang, L.; Xie, X. Q., Modeling, molecular dynamics simulation, and mutation validation for structure of cannabinoid receptor 2 based on known crystal structures of GPCRs. *J Chem Inf Model* **2014,** *54* (9), 2483-2499.
48. Xie, X. Q.; Yang, P.; Zhang, Y.; Zhang, P.; Wang, L.; Ding, Y.; Yang, M.; Tong, Q.; Cheng, H.; Ji, Q.; McGuire, T.; Yuan, W.; Cheng, T.; Gao, Y., Discovery of novel INK4C small-molecule inhibitors to promote human and murine hematopoietic stem cell ex vivo expansion. *Sci Rep* **2015,** *5*, 18115.
49. Ma, C.; Wang, L.; Yang, P.; Myint, K. Z.; Xie, X.-Q., Licabeds Ii. Modeling of Ligand Selectivity for G-Protein-Coupled Cannabinoid Receptors. *J Chem Inf Model* **2013,** *53* (1), 11-26.
50. Ma, C.; Wang, L.; Xie, X.-Q., Ligand classifier of adaptively boosting ensemble decision stumps (licabeds) and its application on modeling ligand functionality for 5HT-subtype GPCR families. *J Chem Inf Model* **2011,** *51* (3), 521-531.
51. Myint, K.; Xie, X.-Q., Fragment-based QSAR algorithm development for compound bioactivity prediction. *Sar Qsar Environ Res* **2011,** *22* (3), 385-410

CORE B (PI: Dr. Ivet Bahar)

- **AIMS**
- **ACCOMPLISHMENTS**
- **RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST**
- **PLAN FOR NEXT PERIOD**
- **PRODUCTS**
  - **PUBLICATIONS**
  - **WEBSITE(S) AND OTHER INTERNET SITE(S)**
  - **TECHNOLOGIES OR TECHNIQUES**
- **BIBLIOGRAPHY**

CORE B

MOLECULAR, CELLULAR AND SYSTEMS BIOLOGY METHODS AND TOOLS FOR COMPUTATIONAL DRUG ABUSE RESEARCH

EXECUTIVE SUMMARY

Ivet Bahar, Ph.D

AIMS

Specific Aim 1. Developing and implementing computational models, methods and software for evaluating and visualizing the functional (or dysfunctional) structures and dynamics of selected human transporters and receptors (and their mutants) involved in DA or dysregulation of excitatory signal transmission.

Specific Aim 2. Facilitating the design of new molecular intervention methods by developing and disseminating software tools for assessing target druggability and identifying sites that alter allosteric signaling and transport properties of DA targets investigated by the FRPs.

Specific Aim 3. Assisting in the conceptualization of novel therapeutic strategies and advancement of DA research at large, by developing and enabling the use of computational tools for quantitative systems pharmacology (QSP).

(Specific Aims remain unchanged.)

ACCOMPLISHMENTS

Major Activities

Funded research projects (FRPs) and other collaborations

(1) In collaboration with the Sorkin lab (PITT; **FRP**), we explored the trimerization of dopamine transporter (DAT) in the absence and presence of a small molecule, AIM-100, as well as its implication on DAT structural dynamics using a combination of computational and biochemical methods, and single-molecule live-cell imaging assays.[1] We performed *in silico* saturation mutagenesis experiments to evaluate the effect of all 19 substitutions at all sequence positions on DAT function using Rhapsody v2.[1]

(2) In a collaborative study with the Wenzel (PITT; **FRP**), Kagan and Bayir labs at the University of Pittsburgh, we characterized the differential dynamics, specificity, and allostery of lipoxygenase family members[2], and unearthed the highly specific mechanisms of catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE.[3]

(3) In collaboration with the Wan (Northwestern) and Xie (Pitt; **Core A**) labs, we developed a small molecule inhibitor, WX2–43, that specifically intercepts the interaction between PRMT5 and KLF4, thereby enhancing KLF4 degradation. [4]

(4) In collaboration with the Chu lab (PITT; **FRP**), we investigated the protein network complex involving PINK1, VCP and PKA which resulted in a manuscript published in *eNeuro*.[5]

(5) In collaboration with the Amara lab (NIMH;**FRP**), we characterized the modulation mechanism of anion channel gating by C-terminal domains in excitatory amino acid transporters (manuscript in preparation).[6]

(6) In collaboration with the Greger lab (MRC, UK), we identified a novel ligand-binding site specific to GluA3 AMPAR N-terminal domain (NTD), resulting from its unique conformational flexibility that we confirmed with crystal structures trapped in vastly different states.[7]

(7) In collaboration with the Khananshvili lab (Tel-Aviv University, Israel), we identified key residues controlling bidirectional ion movements in Na^+/Ca^{2+} exchanger, which resulted in a manuscript published in *Cell Cacium.*[8]

Methodology and technology development

(8) We developed, **SignDy**, an integrated pipeline for evaluating the signature dynamics of protein families based on elastic network models, which led to a major publication in *Mol Biol Evol* (Aim 1).[9]

(9) We developed methodologies to study the shared dynamics of LeuT superfamily members and the allosteric differentiation by structural irregularities and multimerization (Aim 1), published in *Philos Trans R Soc Lond B Biol Sci*.[10]

(10) A new version of **Rhapsody**[11] has been generated (Rhapsody v2) to incorporate features computed from Pfam sequence alignments[12], namely the Shannon entropy and the coevolution propensity using mutual information as a metric.[13-15] (an extension of previous aims)

(11) We designed **Pharmmaker**[16] for building pharmacophore model using outputs of druggability simulations (DruGUI). The pharmacophore models can be used for virtual screening of libraries of small molecules. A strong aspect of the method is that Pharmmaker uses multiple target conformations dependent on the binding poses of probes where they interact during druggability simulations, meaning that the binding score in virtual screening can be evaluated in a more realistic manner. Also, we can have multiple pharmacophore models with different target conformations and probe poses, which can be analyzed statistically (Aim 2).

(12) We developed **Quartata**, an *in silico* chemogenomics methodology and server for linking drugs/chemicals, targets, pathways and GO annotations (manuscript in preparation) (Aim 3).

(13) We developed **SMOKE**[17] for estimating unknown parameters of dynamical models using statistical model checking techniques. It can utilize both quantitative data and qualitative knowledge for calibrating large models with hundreds of unknown parameters. It was originally developed for analyzing ordinary differential equation (ODE) models of biological networks, and currently being generalized to other modeling formalisms including stochastic models, rule-based models, and hybrid models (Aim 3).

Significant Results

Below we present some examples of results and highlights from Core B research progress accomplished during the past funding period.

(1) ***Trimerization of dopamine transporter triggered by AIM-100 binding: Underlying molecular mechanisms and effect of mutations***. The Sorkin lab (PITT; **FRP**) recently found that a furopyrimidine, AIM-100 triggered oligomerization of dopamine transporters (DATs), which promotes endocytosis and thereby may moderate dopaminergic transmission[18]. Despite the significance of these events in mediating cellular responses, the underlying molecular mechanisms remain unclear. In the present study, we determined *in-silico* three structural models, for possible trimerization of DATs, in accord with the versatility of LeuT fold to stabilize dimeric or higher order constructs with a variety of packing geometries. Site-directed mutagenesis was performed to examine the effect of these elements/sites on DAT oligomerization and endocytosis (enhanced by AIM-100), and the experimental data were further interpreted using a novel machine learning classifier[11] for assessing the impact of single amino acid variants (SAVs) (see **Figure 1**). Overall the study suggests the possibility of controlling the effective dopamine transport upon altering the oligomerization state of DAT by small molecular binding, as a possible intervention strategy to modulate dopaminergic signaling.

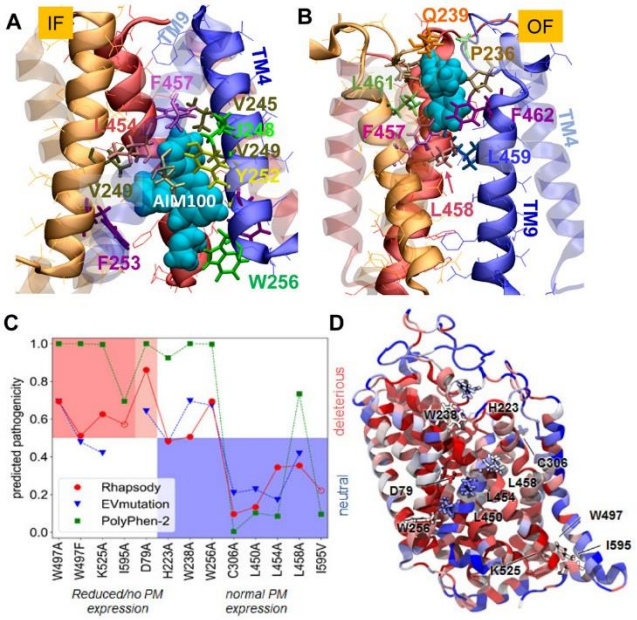


Figure 1: MD simulations reveal the probable binding poses of AIM-100 onto hDAT trimer in the IF (A) and OF (B) state. In silico saturation mutagenesis results and comparison with experimental data. (C) The membrane expression of 13 single amino acid variants (SAVs) has been measured *in vitro*. On the abscissa, these variants are sorted based on the change in their expression levels (*left, red and light-red shade*: no or partial expression; *right, blue shade*: normal/wild-type expression). On the y-axis, the functional impact as predicted by Rhapsody (*red circles*), EVmutation (*blue inverted triangles*) and PolyPhen-2 (*green squares*) is shown. The EVmutation epistatic score has been normalized so that the optimal cutoff between predicted neutral and deleterious effects matches that for Rhapsody and PolyPhen-2 (y = 0.5). Correct predictions lie within the two shaded areas, *red* (true positives) and *blue* (true negatives). **(D)** DAT monomer color-coded by the average pathogenicity shown in **C** (*red/blue*: high/low pathogenicity probability). The 13 mutations sites are labelled and shown in *licorice*.

(2) *Application of Quantitative Systems Pharmacology on drugs of abuse to analyze the molecular mechanisms of drug addiction progress*. We carried out a comprehensive analysis of cellular pathways implicated in a diverse set of 50 drugs of abuse using quantitative systems pharmacology methods (see **Figure 2**). The analysis of the drug/ligand-target interactions compiled in DrugBank and STITCH databases revealed 142 known and 48 newly predicted targets, which have been further analyzed to identify the KEGG pathways enriched at different stages of drug addiction cycle, as well as those implicated in cell signaling and regulation events associated with drug abuse. Apart from synaptic neurotransmission pathways detected as upstream signaling modules that "sense" the early effects of drugs of abuse, pathways involved in neuroplasticity are distinguished as determinants of neuronal morphological changes. Notably, many signaling pathways converge on important targets such as mTORC1. The latter emerges as a universal effector of the persistent restructuring of neurons in response to continued use of drugs of abuse.

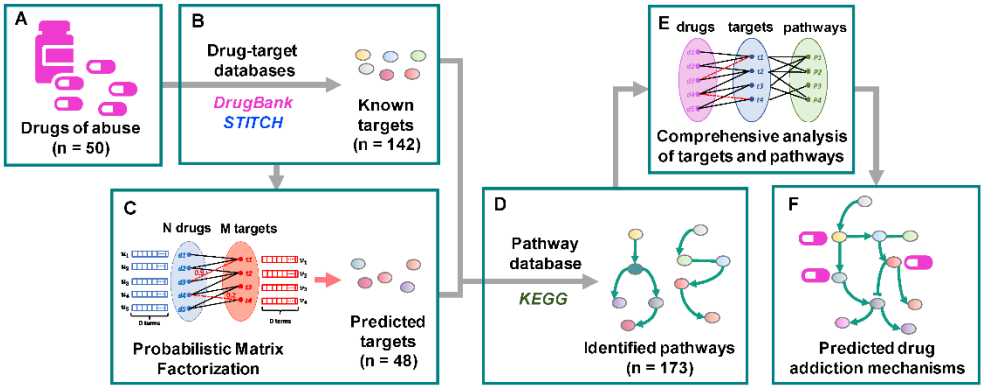


Figure 2. Workflow of the quantitative systems pharmacological analysis. (A) 50 drugs of abuse with a diversity of chemical structures and pharmacological actions were collected as probes. **(B)** 142 known targets of these drugs were identified through drug-target interaction database DrugBank and chemicals interactions database STITCH. **(C)** 48 predicted targets were predicted using our probabilistic matrix factorization (PMF) method[19]. **(D)** 173 human pathways were inferred from the KEGG pathways database by mapping the known and predicted targets. **(E-F)** The pathways were grouped into 5 clusters. The functioning of identified targets and pathways and their involvement in drug addiction were comprehensively examined. More details see ref[20].

(3) *Development of a in silico chemogenomics methodology and server for linking drugs/chemicals, targets, pathways and GO annotations*. We developed an easy and efficient web server QuartataWeb (http://quartata.csb.pitt.edu) (**Figure 3**) for mining known (experimentally verified)

and predicted interactions for 5,494 drugs in DrugBank[21] and 315,514 chemicals in STITCH,[22] along with the confidence levels of the predicted chemical-target interactions (CTIs) using a machine learning based model. In addition, QuartataWeb links targets to KEGG[23] pathways and GO annotations,[24] performs quantitative evaluation of the level of enrichment of pathways and GO annotations given a set of drugs/chemicals or targets. Graphical user interfaces including customized interactive network viewers for CTIs and target-pathway associations facilitate the analysis of outputs and the inference of biological function.

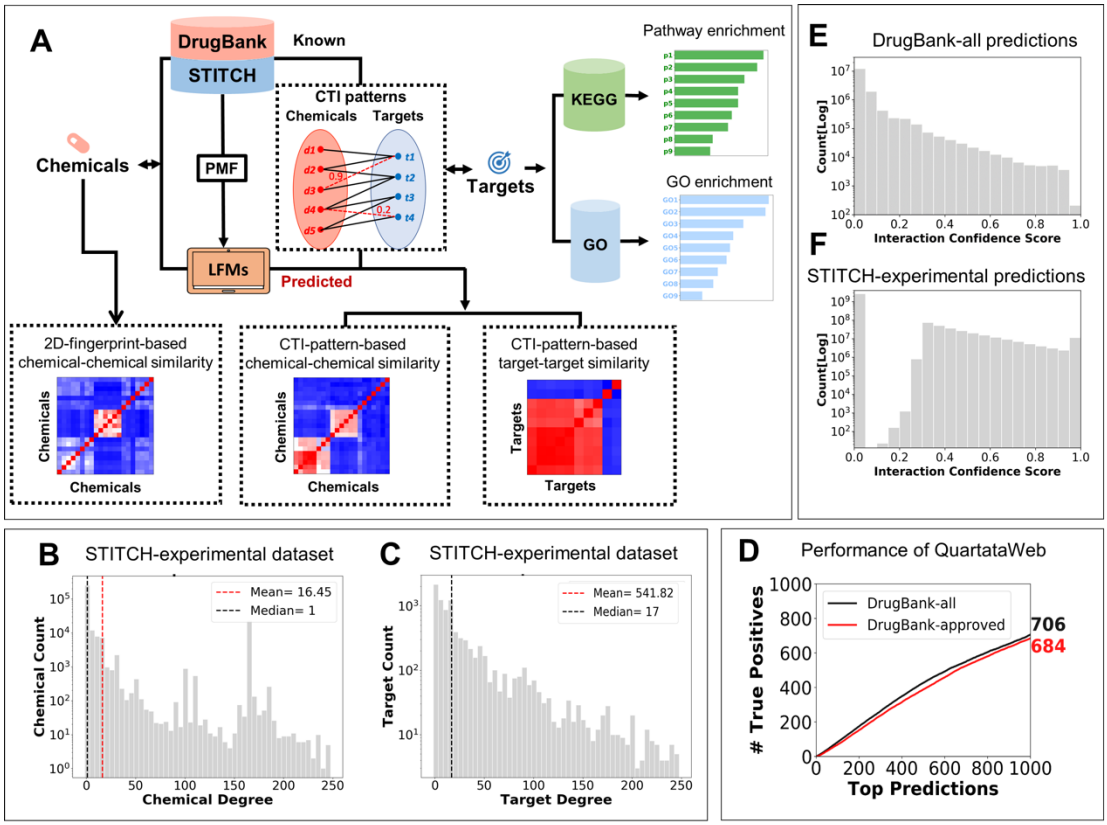


Figure 3. Description of QuartataWeb workflow, data, and performance. A. Schematic description of QuartataWeb workflow**.** Known CTIs extracted from DrugBank and STITCH are used as source datasets for latent factor models (LFMs). The predicted CTIs, chemical-chemical similarity and target-target similarity are pre-computed from the LFMs so as enable easy retrieval. 2D fingerprint similarities between pairs of chemicals are also pre-computed using Python RDKit module (http://rdkit.org). Pathways and GO terms are extracted from KEGG and GO resources, respectively, and stored in our server. The enriched KEGG pathways and GO terms corresponding to the queried targets, or to the known and predicted targets of the queried chemicals, are provided as outputs. **B-C.** Histograms of the degrees of drugs/chemicals and targets in the bipartite CTI networks based on STITCH-experimental dataset. Chemicals and targets exhibit maximal degrees of 4,044, and 26,803, respectively; those with degree less than 250 are displayed here. Dashed vertical lines indicate the mean (*red*) and median (*black*), also written in the inset. **D**. Performance of QuartataWeb shown for DrugBank-approved (*red*) and DrugBank-all (*black*) datasets. The abscissa indicates the rank *m* (1<*m*<1,000) of top-ranking predictions, among all potential CTIs (4,217,124 and 15,408,173 in the respective datasets. The ordinate indicates the average number of recaptured hidden interactions (TPs). For the 1,000 top ranked predictions average precisions f 0.684 and 0.706 are attained in the two respective datasets. **E-F.** Distribution of confidence scores for predicted CTIs computed for DrugBank-all (**E**) and STITCH-experimental (**F**). (submitted)

(4) ***PINK1 interacts with VCP/p97 and activates PKA to promote NSFL1C/p47 phosphorylation and dendritic arborization in neurons*** (eNeuro 2018). While PTEN-induced kinase 1 (PINK1) is well characterized for its role in mitochondrial homeostasis, much less is known concerning its ability to prevent synaptodendritic degeneration. In collaboration with Chu lab (PITT; **FRP**), we found that PINK1 binds and phosphorylates the catalytic subunit of PKA at T197 [PKA_{cat}(pT197)], a site known to activate the PKA holoenzyme. PKA in turn phosphorylates p47 at a novel site (S176) to regulate dendritic complexity. Given that PINK1 physically interacts with both the PKA holoenzyme and the VCP-p47 complex to promote dendritic arborization, we propose that PINK1 scaffolds a novel PINK1-VCP-PKA-p47 signaling pathway to orchestrate dendritogenesis in neurons (**Figure 4**). These findings highlight an important mechanism

by which proteins implicated in Parkinson's disease (PD; PINK1) and frontotemporal dementia (FTD; VCP) interact with each other to support the health and maintenance of neuronal arbors.

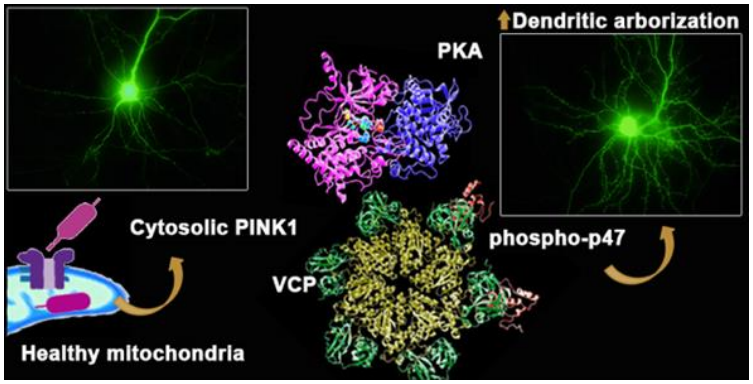


Figure 4: PINK1 interacts with VCP/p97 and activates PKA to promote NSFL1C/p47 phosphorylation and dendritic arborization in neurons. For details see ref[5].

(5) ***Druggability simulations and X-Ray crystallography reveal a ligand-binding site in the GluA3 AMPA receptor N-terminal domain*** (*Structure* 2018) Ionotropic glutamate receptors (iGluRs) mediate the vast majority of excitatory neurotransmission in the brain. Their dysfunction is implicated in several neurological disorders, rendering iGluRs potential drug targets. Here, we performed a systematic analysis of the druggability of two major iGluR subfamilies, using molecular dynamics simulations in the presence of drug-like molecules[7]. We demonstrate the applicability of druggability simulations by faithfully identifying known agonist and modulator sites on AMPA receptors (AMPARs) (see **Figure 5A**) and NMDA receptors. Simulations produced the expected allosteric changes of the AMPAR ligand-binding domain in response to agonist. We also identified a novel ligand-binding site specific to the GluA3 AMPAR N-terminal domain (NTD), resulting from its unique conformational flexibility that we explored further with new crystal structures trapped in vastly different states (see **Figure 5B**). In addition to providing novel insights into iGluR NTD dynamics, our approach identifies druggable sites and permits the determination of pharmacophoric features towards novel iGluR modulators (see **Figure 5C**).

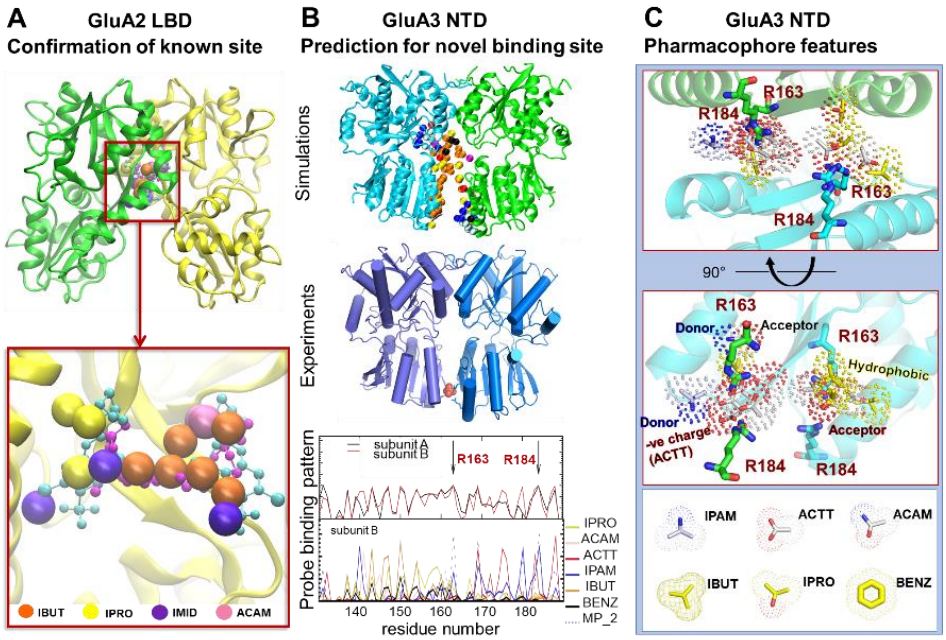


Figure 5. Druggability Simulations, X-Ray Crystallography, and Pharmacophore model for AMPA receptor (A) Druggability MD detecting a known ligand binding site of GluA2 LBD: Large balls are hot spots by probe molecules and their colors are different probes. They closely overlap with the experimentally observed positions of the allosteric modulators cyclothiazide (*cyan balls/stick*; from PDB ID: 1LBC) and (R,R)-2b (*magenta balls/sticks*; PDB ID: 4U5B). **(B)** Novel ligand binding site in GluA3 NTD. (top) Druggability MD probes are shown as spheres near the LL interfaces of the GluA3 NTD dimer. (*middle*) New crystal structure of the GluA3 NTD reveal new dimeric state, which is similar to the open dimer. (bottom)

Number of contacts between probe molecules and residues at the LL interface are observed in the MD simulations. **(C)** Pharmacophore model in GluA3 NTD LL interface. For details see ref.[7]

(6) ***Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity*** (*Mol Biol Evol;* In Press). Recent studies have drawn attention to the evolution of protein dynamics, in addition to sequence and structure, based on the premise structure-encodes-dynamics-encodes-function. Of interest is to understand how functional differentiation is accomplished while maintaining the fold, or how intrinsic dynamics plays out in the evolution of structural variations and functional specificity. We performed a systematic computational analysis of 26,899 proteins belonging to 116 CATH superfamilies. Characterizing cooperative mechanisms and convergent/divergent features that underlie the shared/differentiated dynamics of family members required a methodology that lends itself to efficient analyses of large ensembles of proteins. We therefore introduced, SignDy, an integrated pipeline for evaluating the signature dynamics of families based on elastic network models (**Figure 6**).

Our analysis showed that family members share conserved, highly cooperative (global) modes of motion. Importantly, our analysis discloses a subset of motions that sharply distinguishes subfamilies, which lie in a low-to-intermediate frequency (LTIF) regime of the mode spectrum. This regime has maximal impact on functional differentiation of families into subfamilies, while being evolutionarily conserved among subfamily members. Notably, the high frequency (HF) end of the spectrum also reveals evolutionary conserved features across and within subfamilies; but in sharp contrast to global motions, HF modes are minimally collective. Modulation of robust/conserved global dynamics by LTIF fluctuations thus emerges as a versatile mechanism ensuring the adaptability of selected folds and the specificity of their subfamilies. SignDy further allows for dynamics-based categorization as a new layer of information relevant to distinctive mechanisms of action of subfamilies, beyond sequence or structural classifications.

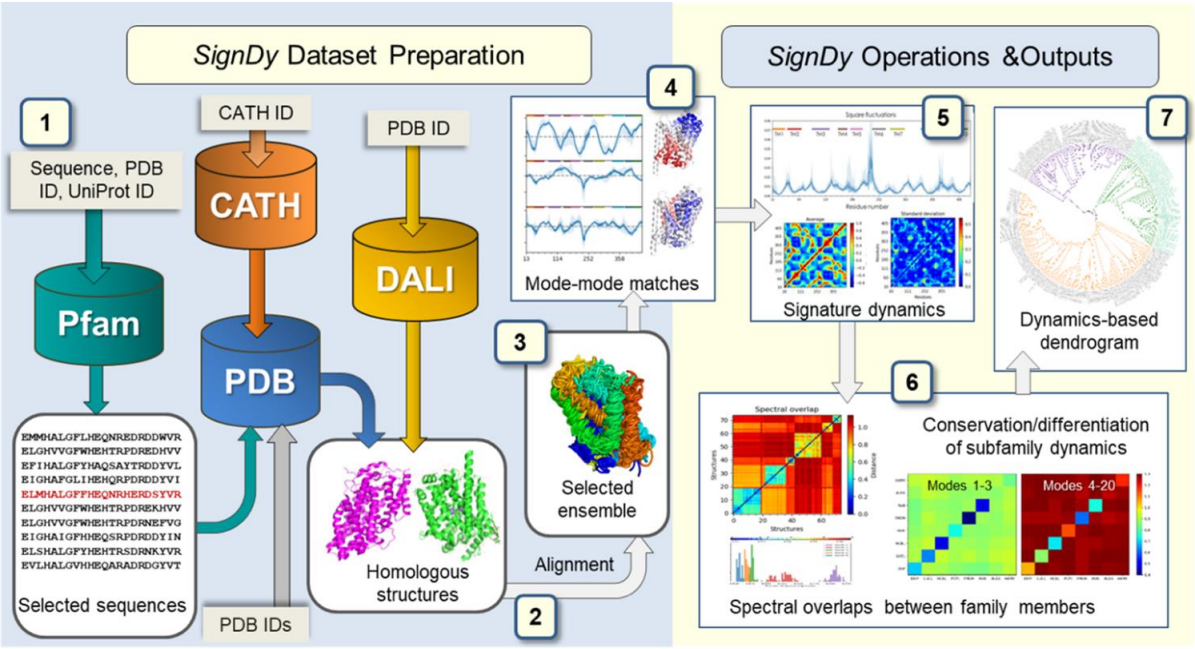


Figure 6: **SignDy workflow.** The workflow is separated into two main parts: dataset preparation (left; steps 1-3) and SignDy operations and outputs (right; steps 4-7). Cylinders and light grey rectangular boxes represent databases and corresponding query inputs, respectively. Details see ref[9].

(7) ***Allosteric differentiation by structural irregularities and multimerization*** (*Philos Trans R Soc Lond B Biol Sci* 2018). The LeuT-fold superfamily includes secondary active transporters from different functional families, which share a common tertiary structure, despite having a remarkably low sequence similarity. By identifying the common structural and dynamical features upon principal component analysis of a comprehensive ensemble of 90 experimentally resolved structures and anisotropic network model evaluation of collective motions, we provide a unified point of view for understanding the reasons why this particular fold has been selected by evolution to accomplish such a broad spectrum of functions. The parallel identification of conserved sequence features, localized at specific sites of transmembrane

helices, sheds light on the role of broken helices (TM1 and TM6 in LeuT) in promoting ion/substrate binding and allosteric interconversion between the outward- and inward-facing conformations of transporters. Finally, the determination of the dynamics landscape for the structural ensemble (**Figure 7**) provides a promising framework for the classification of transporters based on their dynamics, and the characterization of the collective movements that favor multimerization.

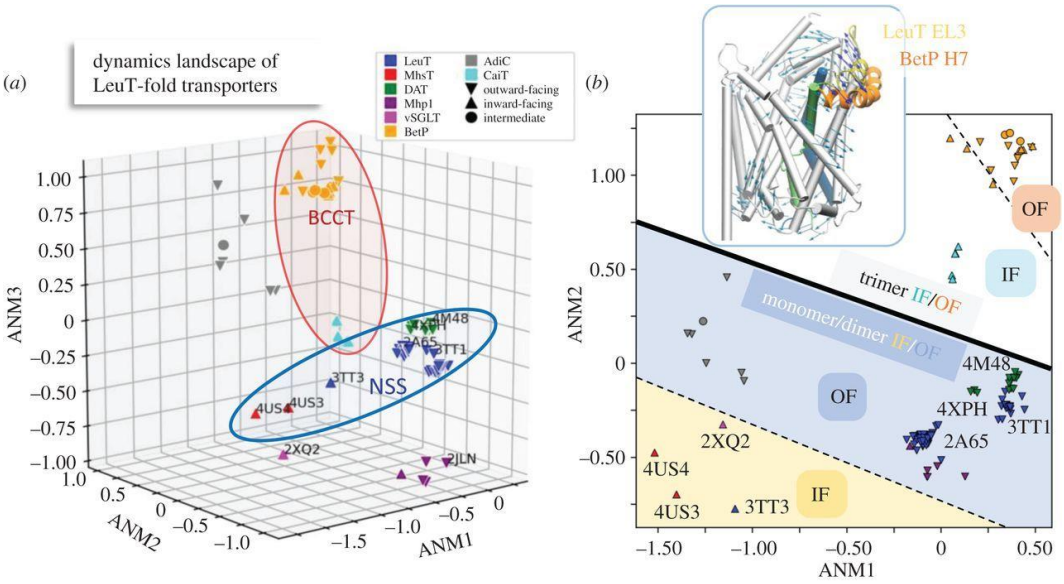


Figure 7. **Classification of LeuT-fold transporters based on their collective motions.** The distribution of the 104 monomers/protomers are displayed in the subspaces of collective modes spanned by the three (a) and two (b) softest ANM modes. Panel (a) shows the clustering of conformers resolved for the same transporter, or those belonging to the same functional families (enclosed in ellipses). Panel (b) provides a clear separation of (i) the monomeric and dimeric transporters (lower left portion; light blue and light yellow) and trimeric transporters (upper right portion), and (ii) the inward-facing and outward-facing conformers within each region. ANM mode 2 (see inset) directs the reconfiguration of the LeuT EL3 (loop–helix, yellow) along a direction (blue arrows on the ribbon diagram) in accord with the structural change undergone by the equivalent BetP H7 helix (orange) upon trimerization. Details see ref[10].

(8) ***Characterization of Differential Dynamics, Specificity, and Allostery of Lipoxygenase Family Members*** (*J. Chem. Inf. Model.* 2019) Accurate modeling of structural dynamics of proteins and their differentiation across different species can help understand generic mechanisms of function shared by family members and the molecular basis of the specificity of individual members. We focused here on the family of lipoxygenases, enzymes that catalyze lipid oxidation, the mammalian and bacterial structures of which have been elucidated. We present a systematic method of approach for characterizing the sequence, structure, dynamics, and allosteric signaling properties of these enzymes using a combination of structure-based models and methods and bioinformatics tools applied to a data set of 88 structures (see **Figure 8**), in collaboration with the Wentzel lab (PITT; **FRP**). The analysis elucidates the signature dynamics of the lipoxygenase family and its differentiation among members, as well as key sites that enable its adaptation to specific substrate binding and allosteric activity.

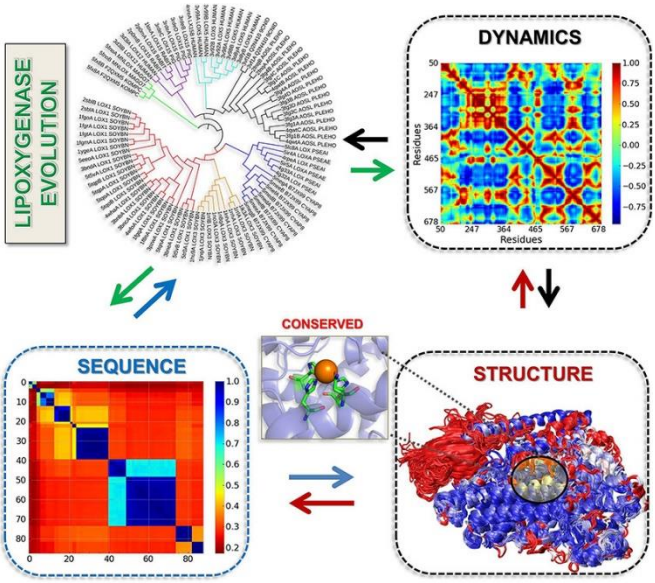


Figure 8: Systematic method of approach for characterizing the sequence, structure, dynamics, and allosteric signaling properties of lipoxygenase family members. Details see Ref[2]

(9) *Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations* (*Protein science* 2019) Pharmmaker is designed for building pharmacophore model using outputs of druggability simulations (DruGUI). The pharmacophore models can be used for virtual screening of libraries of small molecules. This is a suite of programs (see **Figure 9**): (1) druggability molecular dynamics simulations using probe compounds that contain drug-like functional groups and/or features shared by the lead compounds. (2) Identify high affinity residues on the target protein for each probe molecule type from druggability simulations; [25,26] (3) Preselect high affinity residues near a druggable site; (4) Analyze interactions between high affinity residues and probe at the druggable site for each probe type and rank the interaction pairs between residue and probe; (5) Select snapshots with the top ranking interaction pairs; (6) The selected snapshots have target conformations and poses of probes, and these are used for the construction of pharmacophore models, and the pharmacophore models are then used as filters for identifying hits in structure-based virtual screening using Pharmit server [27]. A strong aspect of the method is that Pharmmaker uses multiple target conformations dependent on the binding poses of probes where they interact during druggability simulations. Therefore, the binding score in virtual screening can be more evaluated in a more realistic manner. Also, we can have multiple pharmacophore models with different target conformations and probe poses, which can be analyzed statistically. We expect we can find novel binding pockets and potential compounds, opening new avenues for structure-based design of novel allosteric modulators.

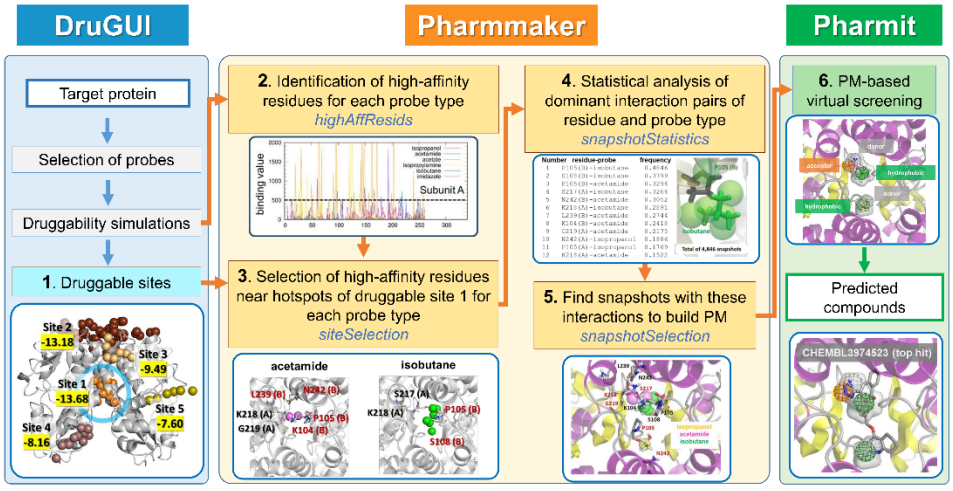


Figure. 9. Computational protocol to select compounds using Pharmmaker. Pharmmaker is for constructing pharmacophore models, in conjunction with druggability simulations (DruGUI) and virtual screening (Pharmit).

Summary of accomplishments

(1) Molecular modeling NSS mechanics led to three publications:

Cheng MH, **Bahar I.** (2019) Monoamine transporters: structure, intrinsic dynamics and allosteric regulation. *Nature Structural & Molecular Biology* 26, 545–556.

Lee JY, Krieger J, Herguedas B, García-Nafría J, Dutta A, Shaikh SA, Greger IH, **Bahar I.** (2019) Druggability Simulations and X-ray Crystallography Reveal a Ligand-binding Site in the GluA3 AMPA Receptor N-terminal Domain. *Structure* 27: 241-252, PMID: 30528594.

Ponzoni L, Zhang S, **Cheng MH**, **Bahar I**. (2018) Shared dynamics of LeuT superfamily members and allosteric differentiation by structural irregularities and multimerization. *Philos Trans R Soc Lond B Biol Sci* 373: 1749 PMID: 29735731

(2) In collaboration with Xie lab (Core A), and Wan lab (P/FP), our computational and experimental studies on small-molecule antagonizes KLF4 methylation has been published on *EBioMedicine:*

Zhou Z, Feng Z, Hu D, Yang P, Gur M, **Bahar I**, Cristofanilli M, Gradishar WJ, **Xie XQ**, **Wan Y**. (2019) A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine* 19: 30312-30313.

(3) In collaboration with Sorkin lab (FRP6), our computational and experimental studies of the trimerization of dopamine transporters has been published on *Neuropharmacology*:

Cheng MH, Ponzoni L, Sorkina T, Lee JY, Zhang S, **Sorkin A**, **Bahar I**. (2019) Trimerization of Dopamine Transporter Triggered by AIM-100 Binding: Molecular Mechanisms and Effect of Mutations. Neuropharmacology [Epub ahead of print] PMID: 31228486

(4) In collaboration with Wenzel lab (FRP13), our computational and experimental studies on lipoxygenases led to two publications:

Anthonymuthu T, Kenny E, Shrivastava I, Tyurina YY, Hier Z, Ting H-C, Dar H, Tyurin V, Nesterova A, Amoscato A, Mikulska-Ruminska K, Rosenbaum J, Mao G, Jinming Z, Conrad M, Kellum J, **Wenzel S**, VanDemark A, **Bahar I**, Kagan V, Bayir H (2018) Empowerment of 15-lipoxygenase catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE. J Am Chem Soc 2018, 140 (51), pp 17835-17839 PMID: 30525572.

Mikulska-Ruminska K, Shrivastava IH, Krieger JM, Zhang S, Li H, Bayir H, **Wenzel SE**, VanDemark AP, Kagan VE, **Bahar I**. (2019) Characterization of differential dynamics, specificity, and allostery of lipoxygenase family members. *J Chem Inf Model*. [Epub ahead of print] PMID: 30762363.

(5) In collaboration with Chu lab (the proposed FRP in the next funding term), our integrated study on NSFL1C/p47 phosphorylation and dendritic arborization in neurons has been published on *eNeuro:*

Wang K, Steer E, Otero PA, Bateman N, **Cheng MH**, Scott A, Wu C, **Bahar I**, Shih Y-T, Hsueh Y-P, **Chu C** (2018) PINK1 Interacts with VCP/p97 and Activates PKA to Promote NSFL1C/p47 Phosphorylation and Dendritic Arborization in Neurons. *eNeuro* 5 (6) ENEURO.0466-18.2018. PMID: 30783609.

(6) Quantitative Systems Pharmacological (QSP) analysis of absued drug associaed targets and pathways led to a publicaiton:

Pei F., Li H., **Liu B**. and **Bahar I.** (2019) Quantitative Systems Pharmacological Analysis of Drugs of Abuse Reveals the Pleiotropy of Their Targets and the Effector Role of mTORC1. *Front. Pharmacol.*, 10, 1-16.

(7) Development of methodologies led to four publications:

Zhang S., Li H., Krieger J., **Bahar I**. Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity. *Mol Biol Evol 36* (9), 2053–2068.

Liu, B.; Gyori, B.; Thiagarajan, P. S. Statistical Model Checking based analysis of biological networks. *Automated Reasoning for Systems Biology and Medicine* 2019; pp 63-92.

Lee, J. Y., Krieger, J., Li, H. & **Bahar, I**. Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations *Protein Science*, in press (2019).

Taylor DL, Gough A, Schurdak ME, Vernetti L, Chennubhotla CS, Lefever D, Pei F, Faeder JR, Lezon TR, Stern AM, **Bahar I**. (2019) "Harnessing Human Microphysiology Systems as Key Experimental Models for Quantitative Systems Pharmacology" in Handbook of Experimental Pharmacology, p 1-41.

OPPORTUNITIES FOR TRAINING & PROFESSIONAL DEVELOPMENT

Mary Hongying Cheng, PhD in Chemical Engineering
Mary's research interest and expertise lie in protein modeling and medicinal chemistry, with focus on molecular mechanism of (i) transporter function, (ii) drug modulation (iii) ion transport through membrane protein channels and (iv) protein-lipid and lipid-lipid interactions. Her current research focuses on understanding neurotransmitter transport mechanisms.

Ji Young Lee, PhD in Physics
Ji Young is specialized in molecular computations for gaining insights into the activation/ inactivation of proteins involved in neurosignaling. He is currently working on ionotropic glutamate receptors ion channels mediating excitatory neurotransmission. Using the Anisotropic Network Model (ANM), he has found that two iGluR families, AMPA receptor (AMPAR) and NMDA receptor (NMDAR), share robust movements. He is actively involved in further development and upgrades of the ProDy API maintained by the Bahar lab.

Bing Liu, PhD in Computer Science
Bing Liu's research area centers on computational systems biology. His work builds mathematical models to describe the dynamics of biological processes. He has developed probabilistic techniques to address stochasticity in biological systems and leveraged machine learning, formal methods, and high-performance computing techniques to enable various analyses of large-scale systems.

HongChun Li, PhD in Chemical Biology
Dr. Li's expertise is the development of the API for QSP; development of server and database for elastic network models (e.g. Gaussian network model database iGNM2.0)**;** protein sequence classification based on machine learning methods, and drug-target association based on probabilistic matrix factorization.

She (John) Zhang, PhD Student, CMU/Pitt Computational Biology
John is currently working on applying elastic network models to several proteins including those sharing the LeuT fold family to determine their principal changes in conformations about the known structures and relate these structural changes to their respective biological functions**.** He is contributing to the development and upgrades of the ProDy API maintained by the Bahar lab.

Fen Pei, PhD Student, CMU/Pitt Computational Biology
Fen is currently working on applying QSP to several proteins including those of abuse drug targeted proteins**.** She is contributing to the development and upgrades of the Balestra Web server.

RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST

(1) All the results are submitted for publication, resulting in 16 peer-reviewed journal publications/book chapters (Jul 2018-Sep 2019). More details see **C.1. Publications.**

(2) The codes are made accessible online, and two servers, **ProDy** and **Balestra**, are being further developed to facilitate the broader use by the community. More details see **C.3. Technologies or techniques.** Furthermore, we have implemented

(3) Presentations and posters were presented in annual society meetings

1. Ji Young Lee, James Krieger, Beatriz Herguedas, Javier García-Nafría, Anindita Dutta, Saher A. Shaikh, Ingo H Greger, and Ivet Bahar, “A ligand-binding site in the GluA3 AMPA receptor N-terminal domain observed in druggability simulations and X-ray crystallography”, Biophysical Society 63th Annual Meeting, Baltimore, Maryland, USA, Mar 2019
2. James M. Krieger, Béatriz Herguedas, Bishal Singh, Jiyoung Lee, Burak Kaynak, Ingo Greger, and Ivet Bahar, “Dynamics of AMPA Receptors from Simulations and Electron Microscopy” Biophysical Society 63th Annual Meeting, Baltimore, Maryland, USA, Mar 2019; Biophysical Journal, Vol. 116, Issue 3, p344a Published in issue: February 15, 2019
3. Luca Ponzoni and Ivet Bahar “Structural Dynamics is a Determinant of the Functional Significance of Missense Variants”, Workshop at Temple University, Philadelphia: "Statistical Mechanics of Protein Sequences: from fitness to free energy landscapes and back", May 4th 2018.
4. She Zhang and Ivet Bahar “Cell-to-Cell Heterogeneities at Genome Scale Investigated by the Gaussian Network Model”, Pitt Day in Harrisburg, May 23, 2018.
5. Bing Liu. “Systems Biology, Artificial Intelligence, and Better Life”, Fudan-Guanghua International Forum for Young Scholars, Shanghai, China, December 2018.

(4) Invited Talks (Core B PI Ivet Bahar)

1. Invited talk by Bahar, at Jacques Monod Conference, Sciences biologiques Ecologie et Environnement, Ligand-gated ion channels from atomic structure to synaptic transmission. Roscoff, France, on May 20-24, 2019:
2. Seminar by I. Bahar. National Institute of Health (NIH) National Institute on Drug Abuse (NIDA) Intramural Research Program (IRP) Seminar Series. Baltimore, MD, USA (May 14, 2019)
3. Invited Speaker (I. Bahar). Multiscale Modeling of Chromatin: Bridging Experiment with Theory. (Biophysical Society Thematic Meeting.) Les Houches, France. March 31-April 5, 2019
4. Invited Symposium Speaker (I. Bahar) at 63rd Annual Meeting of the Biophysical Society, Symposium on Glutamate Receptors. Baltimore, MD, March 2-4, 2019
5. Invited Speaker (I. Bahar) at Conference on Modeling of Protein Interactions (MPI). Lawrence, Kansas .Nov 8-10, 2018
6. Invited Speaker (I. Bahar) at CECAM (Centre European pour le Calcul Atomique et Moléculaire) Workshop, "Multiscale simulations of allosteric regulatory mechanisms in cancer-associated proteins and signaling protein networks," Lugano, Switzerland Oct 15-17, 2018.
7. Invited Speaker (I. Bahar). Computational Biology Workshop, Arizona State University, Phoenix, Arizona. Oct 6-8, 2018
8. Invited Speaker (I. Bahar) Inaugural International Transmembrane Transporter Society (ITTS) Symposium. Vienna, Austria. Session Organizer and Speaker (Sept 18-21, 2018)
9. Invited Speaker (I. Bahar) CECAM (Centre European pour le Calcul Atomique et Moléculaire) Workshop, "Normal modes of biological macromolecules: methods and applications," Paris, France (Sept 12-14, 2018)
10. Session Organizer (I. Bahar) European Conference in Computational Biology (ECCB) 2018, jointly held with the ISMB (International Society for Computational Biology). Athens, Greece. (Sept 9-12, 2018).

11. Invited Speaker (I. Bahar) 6th Annual iGluRetreat, jointly held with the Department of Chemistry of Carnegie Mellon University and the Department of Neuroscience of the University of Pittsburgh. Pittsburgh, PA. July 31-August 2, 2018.
12. 6th Annual iGluRetreat, jointly held with the Department of Chemistry of Carnegie Mellon University and the Department of Neuroscience of the University of Pittsburgh. Pittsburgh, PA. Invited Speaker.

PLANS FOR NEXT PERIOD

Below is a summary of anticipated Core B activities for the coming year:

(1) Coarse-grained molecular dynamics (CGMD) simulations of complexes formed by DAT and SERT (Aim 1). One of our current main interests is to understand interaction between DATs and membrane lipids. Some our ongoing studies are focusing on this issue. Our current CGMD results show that DAT variations and membrane are closely related and furthermore DAT clustering is also regulated by membrane. We plan to investigate in more details on the function of DAT, its clustering and interaction with neuronal lipids using CGMD. We will extend to SERT-SERT complex, taking advantage of the newly resolved SERT structures in multiple conformations.

(2) Improvement of MCell simulations of dopamine signaling (an extension of previous aims) We recently developed a multiscale model for examining the effects of spatial complexity and firing patterns on dopamine reuptake[28]. We used electron microscopy (EM) images and immunofluorescence of transgenic knock-in mouse brains that express HA-tagged DAT in dopamine neurons (performed in Sorkin lab; DBP3) to construct a realistic *MCell* model[28]. DAT spatial distributions and structural heterogeneities were observed to alter the efficiency of DAT, demonstrating that realistic spatial descriptions are required to accurately simulate DA reuptake. We plan to improve our current MCell modeling by implementing new features: (1) *2D diffusion of DAT*. Our previous simulations did not include the effect of the diffusion of DAT on the membrane of the synaptic cells. We will upgrade *MCell* to allow for the diffusion of membrane proteins, and examine the effect of their clustering or oligomerization on dopamine reuptake and signaling properties by conducting simulations that are at least one order of magnitude longer than our previous (100 milliseconds) simulations.(2) There are conflict results reported in the literature regarding the lipid microdomain localization, trafficking and regulation of dopamine transporters and receptors (see recent review[29]). Different pathways and trafficking may exist for the “raft” and “non-raft” DATs[30,31], to which DAT is dynamically sequestrated[32]. Non-raft populations may be responsible for AMPH-induced DAT internalization, whereas raft populations may dominate DAT-mediated efflux[30].We plan to improve our current MCell simulation to enable subcellular distribution/simulations of DATs in the lipid microdomains.

(3) Further development of Rhapsody (an extension of previous aims). The general idea behind the 2018 implementation of the algorithm[11] is maintained in the updated version. Namely, we propose an approach for predicting the outcome of amino acid substitutions in (human) proteins, based on a random forest classifier trained over not only well-established sequence-based (e.g. conservation of wild-type/mutated amino acid) and structural (computed from PDB structures) properties, but also *dynamical* properties, derived from coarse-grained Elastic Network Models (ENM). The focus on the dynamical characterization of mutation sites provides an orthogonal approach to the problem of variants’ classification that allows us to stand out with respect to other analogous methods. In the second iteration of the algorithm (RHAPSODY), we pursue the goal of a better integration between different computational strategies, by collecting a more diverse set of features. Major upgrades will include: (i) an expanded feature set, now comprising conservation and *coevolution* properties extracted from Pfam domains, (ii) a more refined dataset of about 20,000 human missense variants, built from consensus between clinical interpretations of variants found in multiple databases, (iii) the flexibility of using a custom PDB structure for the evaluation of structural and dynamical features, such as specific conformational/oligomerization states, homologous structures from other organisms and theoretical homology models. Finally, the Python code has been completely overhauled and integrated with the ProDy library developed in our group, and a new, more user-friendly webserver will be made available to the public.

(4) Exploring conformational landscapes with collective molecular dynamics (CoMD) and comparison to MDeNM and ClustENM (Aim 3). Our lab groups recently developed methods that use normal mode analysis together with molecular dynamics simulations to understanding protein collective motions. These methods were developed with different goals in mind and have not been systematically compared. Collective molecular dynamics (CoMD) also has not been used much for exploring conformational landscapes. We therefore developed CoMD to this aim and compared it against the other two methods, MDeNM[33] and ClustENM[34]. We used proteins of various sizes as test cases including HIV protease (~ 200 residues) and triosphosphate isomerase (~ 500 residues).

(5) Further development and implementation of QuartataWeb (Aim 3)***.*** Building on our original work for mining protein-drug interactions using DrugBank as a data source, we have now extended the methodology (probabilistic matrix factorization [19]) to be able to retrieve and analyze data from a significantly larger database, STITCH,[35] of protein-small molecule interactions. The new resource, called QuartataWeb ((http://quartata.csb.pitt.edug), is an advanced version of the server BalestraWeb. [36] It now permits us to analyze the updated versions of DrugBank and STITCH, in addition to linking drugs and targets to pathways, using KEGG Pathways database. We anticipate QuartataWeb to serve as an important resource for enabling *in silico* chemogenomics analysis and linking drugs/chemicals, targets, pathways and GO annotations.

PRODUCTS

Publications:

Sixteen publications from Core B acknowledged the P30 grant (Jul 2018-Sep 2019):

1. **Cheng MH**, **Bahar I.** (2019) Monoamine transporters: structure, intrinsic dynamics and allosteric regulation. *Nature Structural & Molecular Biology* 26, 545–556.

2. Zhou Z, Feng Z, Hu D, Yang P, Gur M, **Bahar I**, Cristofanilli M, Gradishar WJ, **Xie XQ**, Wan Y. (2019) A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine* 19: 30312-30313.

3. Taylor DL, Gough A, Schurdak ME, Vernetti L, Chennubhotla CS, Lefever D, Pei F, Faeder JR, Lezon TR, Stern AM, **Bahar I**. (2019) "Harnessing Human Microphysiology Systems as Key Experimental Models for Quantitative Systems Pharmacology" in Handbook of Experimental Pharmacology, p 1-41.

4. Thermozier S, Zhang X, Hou W, Fisher R, Epperly MW, **Liu B**, **Bahar I**, Wang H, Greenberger JS. (2019) Radioresistance of Serpinb3a-/- Mice and Derived Hematopoietic and Marrow Stromal Cell Lines. Radiation Research 192, 267-281 PMID: 31295086.

5. **Cheng MH**, Ponzoni L, Sorkina T, Lee JY, Zhang S, **Sorkin A**, **Bahar I**. (2019) Trimerization of Dopamine Transporter Triggered by AIM-100 Binding: Molecular Mechanisms and Effect of Mutations. Neuropharmacology [Epub ahead of print] PMID: 31228486

6. Zhang S., Li H., Krieger J., **Bahar I**. Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity. *Mol Biol Evol 36* (9), 2053–2068.
7. Pei F., Li H., **Liu B**. and **Bahar I.** (2019) Quantitative Systems Pharmacological Analysis of Drugs of Abuse Reveals the Pleiotropy of Their Targets and the Effector Role of mTORC1. *Front. Pharmacol.*, 10, 1-16.

8. Mikulska-Ruminska K, Shrivastava IH, Krieger JM, Zhang S, Li H, Bayir H, **Wenzel SE**, VanDemark AP, Kagan VE, **Bahar I**. (2019) Characterization of differential dynamics, specificity, and allostery of lipoxygenase family members. *J Chem Inf Model*. [Epub ahead of print] PMID: 30762363.

9. Lee JY, Krieger J, Herguedas B, García-Nafría J, Dutta A, Shaikh SA, Greger IH, **Bahar I.** (2019) Druggability Simulations and X-ray Crystallography Reveal a Ligand-binding Site in the GluA3 AMPA Receptor N-terminal Domain. *Structure* 27: 241-252, PMID: 30528594.

10. Wang K, Steer E, Otero PA, Bateman N, **Cheng MH**, Scott A, Wu C, **Bahar I**, Shih Y-T, Hsueh Y-P, **Chu C** (2018) PINK1 Interacts with VCP/p97 and Activates PKA to Promote NSFL1C/p47 Phosphorylation and Dendritic Arborization in Neurons. *eNeuro* 5 (6) ENEURO.0466-18.2018. PMID: 30783609.

11. Van Dijk L, Giladi M, Refaeli B, Hiller R, **Cheng MH**, **Bahar I**, Khananshvili D. (2018) Key residues controlling bidirectional ion movements in Na+/Ca2+ exchanger. *Cell Calcium,* 76: 10-22. PMID: 30248574.

12. Ponzoni L, Zhang S, **Cheng MH**, **Bahar I**. (2018) Shared dynamics of LeuT superfamily members and allosteric differentiation by structural irregularities and multimerization. *Philos Trans R Soc Lond B Biol Sci* 373: 1749 PMID: 29735731

13. Anthonymuthu T, Kenny E, Shrivastava I, Tyurina YY, Hier Z, Ting H-C, Dar H, Tyurin V, Nesterova A, Amoscato A, Mikulska-Ruminska K, Rosenbaum J, Mao G, Jinming Z, Conrad M, Kellum J, **Wenzel S**, VanDemark A, **Bahar I**, Kagan V, Bayir H (2018) Empowerment of 15-lipoxygenase catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE. J Am Chem Soc 2018, 140 (51), pp 17835-17839 PMID: 30525572.

14. **Liu, B**.; Gyori, B.; Thiagarajan, P. S. Statistical Model Checking based analysis of biological networks. *Automated Reasoning for Systems Biology and Medicine* 2019; pp 63-92.

15. Zhang, B.; Li, F.; Chen, Z.; Shrivastava, I. H.; Gasanoff, E. S.; Dagda, R. K. Naja mossambica mossambica cobra cardiotoxin targets mitochondria to disrupt mitochondrial membrane structure and function. *Toxins* 2019**,** *11* (3), 152

16. Lee, J. Y., Krieger, J., Li, H. & **Bahar, I**. Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations *Protein Science*, in press (2019).

Website(s) or other internet site(s):

(1) http://www.ccbb.pitt.edu/Faculty/bahar
Description: main home page

(2) http://prody.csb.pitt.edu/
Description: protein dynamics and sequence analysis: including seven modules

(3) http://ignm.ccbb.pitt.edu/ and http://gnm.csb.pitt.edu/
Description: GNM online servers

(4) http://anm.csb.pitt.edu/cgi-bin/anm2/anm2.cgi
Description: ANM online server

(5) http://balestra.csb.pitt.edu/
Descritption: BalestraWeb addresses drug-protein interactions for repurposable drugs.

(6) http://prody.csb.pitt.edu/drugui/
Descritption: setup and analysis of druggability simulations

Three additional new web servers are being implemented currently and a first alpha version of each has been completed.

(7) http://rhapsody.csb.pitt.edu/

Description: pathogenicity prediction of missense variants by taking structural dynamics into considerations (version 1)

(8) http://quartata.csb.pitt.edu
Description: in silico chemogenomics methodology and server for linking drugs/chemicals, targets, pathways and GO annotations

(9) http://quartata.csb.pitt.edu/amdb
Description: A database and server for searching autophagy modulating drugs/chemicals, targets, pathways, in vitro/in vivo experiments, and clinical trials.

Technologies or techniques:

1. We designed a new module **Pharmmaker**[16] (http://prody.csb.pitt.edu/pharmmaker/) for building pharmacophore model using outputs of druggability simulations (DruGUI). The pharmacophore models can be used for virtual screening of libraries of small molecules. A strong aspect of the method is that Pharmmaker uses multiple target conformations dependent on the binding poses of probes where they interact during druggability simulations, meaning that the binding score in virtual screening can be evaluated in a more realistic manner. Also, we can have multiple pharmacophore models with different target conformations and probe poses, which can be analyzed statistically.
2. We Implemented a new module ***SignDy***[9] to ***ProDy*** *API.* ***SignDy*** calculates the signature dynamics of families of proteins that share similar folds, but not necessarily similar sequences. Signature dynamics includes shared mode profiles, shared covariance between residue fluctuations, and their variations across family members. Additional information can be found in online tutorials;(http://prody.csb.pitt.edu/signdy/).
3. We developed an easy and efficient web server ***QuartataWeb*** (http://quartata.csb.pitt.edu) for mining known (experimentally verified) and predicted interactions for 5,494 drugs in DrugBank[21] and 315,514 chemicals in STITCH,[22] along with the confidence levels of the predicted chemical-target interactions (CTIs) using a machine learning based model.
4. We initiated the implementation of ***Rhapsody*** (http://rhapsody.csb.pitt.edu/) for upgraded pathogenicity prediction of missense variants by taking structural dynamics into considerations.
5. significantly advanced the capabilities of ***ProDy***, which currently offers 10+ modules with user-friendly visualization tools, more than 40,000 code-line, and more than 4,000 pages of documentation including manuals and tutorial. *ProDy*[13,14] reached an impressive milestone of 2 million downloads (http://prody.csb.pitt.edu/statistics/) as of September 2019.
6. Our database of GNM results, **iGNM DB**[15] **(***Nucleic Acids Res* **44**: D415-422; 2016) now covers 95% of structures available in the PDB (a 5-fold increase compared to earlier version); and its improved techniques, libraries and markup language (Ajax, JQuery, HTM5, PHP and Highcharts) enhanced its security and interoperability.
7. Our webserver **DynOmics**[16] (dynomics.pitt.edu) **(***Nucleic Acids Res* 2017) is a portal developed to leverage rapidly growing structural proteomics data by efficiently and accurately evaluating the dynamics of structurally resolved systems, from individual molecules to large complexes and assemblies, in the context of their physiological environment.
8. Updated web server **Balestra**[17] **(**http://balestra.csb.pitt.edu/) (*Bioinformatics* **31**:131-3**;** 2015) using the DrugBank version 5. Database architecture (PostgreSQL) has been used for the BalestraWeb server to improve the query performance. The searching engine has been improved to accept multiple proteins/drugs. We have integrated the protein information from the Uniprot database. We have improved the GUI/interface of the input and the output with the integrated information of proteins and drugs. The server has been extended to the Stitch database version 4 using the PostgreSQL database. We are developing the new version of BalestraWeb, which can efficiently identify chemicals, targets and pathways for drug abuse and will play an important role in discovering the underlying mechanisms and developing corresponding therapeutic strategies.
9. Implemented ***DruGUI*** (http://prody.csb.pitt.edu/drugui/) as a VMD plugin designed for setup and analysis of simulations containing small organic molecules (probes) for druggability assessment. *DruGUI* can incorporate a diverse set of molecules from CHARMM General Force Field (CGenFF) into simulations. *DruGUI* is used to i) identify *druggable* and *ligandable* sites; ii) setup a simulations that contain diverse probe molecules; iii) calculate probe molecule occupancy

grids; iv) analyze druggability of target protein; and v) perform druggability analysis of specific sites.

10. Our **SMOKE**[17] (https://liubing1020.github.io/smoke/) is a Statistical MOdel checKing tool for Estimating unknown parameters of dynamical models. It can utilize both quantitative data and qualitative knowledge for calibrating large models with hundreds of unknown parameters. It was originally developed for analyzing ordinary differential equation (ODE) models of biological networks, and currently being generalized to other modeling formalisms including stochastic models, rule-based models, and hybrid models.

References

1. Cheng, M.H., Ponzoni, L., Sorkina, T., Lee, J.Y., Zhang, S., Sorkin, A. & Bahar, I. Trimerization of dopamine transporter triggered by AIM-100 binding: Underlying molecular mechanisms and effect of mutations. *Neuropharmocology* **In press**(2019).
2. Mikulska-Ruminska, K., Shrivastava, I., Krieger, J., Zhang, S., Li, H., Bayır, H., Wenzel, S.E., VanDemark, A.P., Kagan, V.E. & Bahar, I. Characterization of Differential Dynamics, Specificity, and Allostery of Lipoxygenase Family Members. *J. Chem. Inf. Model* **59**, 2496-2508 (2019).
3. Anthonymuthu, T.S., Kenny, E.M., Shrivastava, I., Tyurina, Y.Y., Hier, Z.E., Ting, H.-C., Dar, H.H., Tyurin, V.A., Nesterova, A. & Amoscato, A.A. Empowerment of 15-lipoxygenase catalytic competence in selective oxidation of membrane ETE-PE to ferroptotic death signals, HpETE-PE. *Journal of the American Chemical Society* **140**, 17835-17839 (2018).
4. Zhou, Z., Feng, Z., Hu, D., Yang, P., Gur, M., Bahar, I., Cristofanilli, M., Gradishar, W.J., Xie, X.-q. & Wan, Y. A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine* (2019).
5. Wang, K.Z.Q., Steer, E., Otero, P.A., Bateman, N.W., Cheng, M.H., Scott, A.L., Wu, C., Bahar, I., Shih, Y.-T., Hsueh, Y.-P. & Chu, C.T. PINK1 Interacts with VCP/p97 and Activates PKA to Promote NSFL1C/p47 Phosphorylation and Dendritic Arborization in Neurons. *eneuro* **5**, ENEURO.0466-0418.2018 (2018).
6. Neureiter, E., Torres-Salazar, D., Cheng, M., Kara, A., Gonzalez-Suarez, A., Glasser, C., Garcia-Olivares, J., Bahar, I. & Amara, S. Modulation of Anion Channel Gating by C-terminal Domains in Excitatory Amino Acid Transporters. *Society for Neuroscience 2019 Meeting* (2019).
7. Lee, J.Y., Krieger, J., Herguedas, B., García-Nafría, J., Dutta, A., Shaikh, S.A., Greger, I.H. & Bahar, I. Druggability Simulations and X-Ray Crystallography Reveal a Ligand-Binding Site in the GluA3 AMPA Receptor N-Terminal Domain. *Structure* **27**, 241-252. e243 (2019).
8. van Dijk, L., Giladi, M., Refaeli, B., Hiller, R., Cheng, M.H., Bahar, I. & Khananshvili, D. Key residues controlling bidirectional ion movements in Na+/Ca2+ exchanger. *Cell calcium* **76**, 10-22 (2018).
9. Zhang, S., Li, H., Krieger, J. & Bahar, I. Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity. *Mol Biol Evol* **36** (9), 2053–2068. (2019).
10. Ponzoni, L., Zhang, S., Cheng, M.H. & Bahar, I. Shared dynamics of LeuT superfamily members and allosteric differentiation by structural irregularities and multimerization. *Philos Trans R Soc Lond B Biol Sci* **373**, 20170177 (2018).
11. Ponzoni, L. & Bahar, I. Structural dynamics is a determinant of the functional significance of missense variants. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4164-4169 (2018).
12. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E. & Finn, R.D. The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427-D432 (2018).
13. Bakan, A., Dutta, A., Mao, W., Liu, Y., Chennubhotla, C., Lezon, T.R. & Bahar, I. Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* **30**, 2681-2683 (2014).
14. Dunn, S.D., Wahl, L.M. & Gloor, G.B. Mutual information without the influence of phylogeny or entropy dramatically improves residue contact prediction. *Bioinformatics* **24**, 333-340 (2008).
15. Gloor, G.B., Martin, L.C., Wahl, L.M. & Dunn, S.D. Mutual information in protein multiple sequence alignments reveals two classes of coevolving positions. *Biochemistry* **44**, 7156-7165 (2005).
16. Lee, J.Y., Krieger, J., Li, H. & Bahar, I. Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations *Protein Science*, in press (2019).

17. Liu, B., Gyori, B. & Thiagarajan, P. Statistical model checking based analysis of biological betworks. *Automated Reasoning for Systems Biology and Medicine*, 63-92 (2019).
18. Sorkina, T., Ma, S., Larsen, M.B., Watkins, S.C. & Sorkin, A. Small molecule induced oligomerization, clustering and clathrin-independent endocytosis of the dopamine transporter. *eLife* **7**, e32293 (2018).
19. Cobanoglu, M.C., Liu, C., Hu, F., Oltvai, Z.N. & Bahar, I. Predicting drug-target interactions using probabilistic matrix factorization. *J Chem Inf Model* **53**, 3399-3409 (2013).
20. Pei, F., Li, H., Liu, B. & Bahar, I. Quantitative Systems Pharmacological Analysis of Drugs of Abuse Reveals the Pleiotropy of Their Targets and the Effector Role of mTORC1. *Frontiers in Pharmacology* **10**, 1-16 (2019).
21. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C. & Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074-D1082 (2018).
22. Szklarczyk, D., Santos, A., von Mering, C., Jensen, L.J., Bork, P. & Kuhn, M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* **44**, D380-384 (2016).
23. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361 (2017).
24. Huntley, R.P., Sawford, T., Mutowo-Meullenet, P., Shypitsyna, A., Bonilla, C., Martin, M.J. & O'Donovan, C. The GOA database: gene Ontology annotation updates for 2015. *Nucleic Acids Res* **43**, D1057-D1063 (2015).
25. Bakan, A., Kapralov, A.A., Bayir, H., Hu, F., Kagan, V.E. & Bahar, I. Inhibition of peroxidase activity of cytochrome c: De novo compound discovery and validation. *Molecular pharmacology* **88**, 421-427 (2015).
26. Bakan, A., Nevins, N., Lakdawala, A.S. & Bahar, I. Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules. *Journal of chemical theory and computation* **8**, 2435-2447 (2012).
27. Sunseri, J. & Koes, D.R. Pharmit: interactive exploration of chemical space. *Nucleic acids research* **44**, W442-W448 (2016).
28. Kaya, C., Cheng, M.H., Block, E.R., Bartol, T.M., Sejnowski, T.J., Sorkin, A., Faeder, J.R. & Bahar, I. Heterogeneities in Axonal Structure and Transporter Distribution Lower Dopamine Reuptake Efficiency. *eNeuro* **5**, ENEURO.0298-0217.2017 (2018).
29. Liu, J.J., Hezghia, A., Shaikh, S.R., Cenido, J.F., Stark, R.E., Mann, J.J. & Sublette, M.E. Regulation of monoamine transporters and receptors by lipid microdomains: implications for depression. *Neuropsychopharmacology* **43**, 2165-2179 (2018).
30. Robertson, S., Matthies, H. & Galli, A. A closer look at amphetamine-induced reverse transport and trafficking of the dopamine and norepinephrine transporters. *Molecular neurobiology* **39**, 73-80 (2009).
31. Foster, J.D., Adkins, S.D., Lever, J.R. & Vaughan, R.A. Phorbol ester induced trafficking-independent regulation and enhanced phosphorylation of the dopamine transporter associated with membrane rafts and cholesterol. *Journal of neurochemistry* **105**, 1683-1699 (2008).
32. Rahbek-Clemmensen, T., Lycas, M.D., Erlendsson, S., Eriksen, J., Apuschkin, M., Vilhardt, F., Jørgensen, T.N., Hansen, F.H. & Gether, U. Super-resolution microscopy reveals functional organization of dopamine transporters into cholesterol and neuronal activity-dependent nanodomains. *Nat Commun* **8**, 740 (2017).
33. Costa, M.G., Batista, P.R., Bisch, P.M. & Perahia, D. Exploring free energy landscapes of large conformational changes: molecular dynamics with excited normal modes. *Journal of chemical theory and computation* **11**, 2755-2767 (2015).
34. Kurkcuoglu, Z., Bahar, I. & Doruker, P. ClustENM: ENM-based sampling of essential conformational space at full atomic resolution. *Journal of chemical theory and computation* **12**, 4549-4562 (2016).
35. Kuhn, M., Szklarczyk, D., Franceschini, A., Von Mering, C., Jensen, L.J. & Bork, P. STITCH 3: zooming in on protein–chemical interactions. *Nucleic acids research* **40**, D876-D880 (2011).
36. Cobanoglu, M.C., Oltvai, Z.N., Taylor, D.L. & Bahar, I. BalestraWeb: efficient online evaluation of drug–target interactions. *Bioinformatics* **31**, 131-133 (2014).

CORE C (PIs: Drs. Eric Xing and Wei Wu)

- **AIMS**
- **ACCOMPLISHMENTS**
- **RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST**
- **PLANS FOR NEXT REPORTING PERIOD**
- **PRODUCTS**
  - **PUBLICATIONS**
  - **TECHNOLOGIES OR TECHNIQUES**
- **BIBLIOGRAPHY**

CORE C

COMPUTATIONAL GENOMICS FOR DRUG ABUSE (CG4DA)

EXECUTIVE SUMMARY

Eric Xing, PhD and Wei Wu PhD

AIMS

Specific Aim 1. To Develop Machine Learning Models and Algorithms for Transcriptome-Wide Screening of Expression Traits and Molecular Markers for Drug-Abuse.

Specific Aim 2. To Conduct Genome-Wide Discovery of Drug Targets and Their Epistatic Genetic Influences via Structured Association Mapping.

Specific Aim 3. To Develop Software Systems to Facilitate Drug-Abuse Diagnosis, Mechanistic Research, and Possibly Guide DA Treatment.

(Specific Aims remain unchanged.)

ACCOMPLISHMENTS

(a) We have developed new machine learning tools for DA research (***Aims 1–3***).

i) Many genetic variants have individually smaller effects, but collectively large effects, on complex human diseases. These variants are difficult to discover using conventional statistical methods. In order to discover such variants associated with alcoholism and Alzheimer's disease, we developed a novel machine learning method called Constrained Sparse Linear Mixed Model (CS-LMM) (Aims 1 & 2). Using CS-LMM, we identified multiple potential weak but significant SNP variants associated with alcoholism and Alzheimer's disease (AD).

ii) Motivated by the results we observed from the CS-LMM project that some genes are associated with both alcoholism and AD), we extended the study to develop a new machine learning method Coupled Mixed Model (CMM) that can identify genes that are jointly associated with two different types of diseases (Aims 1 & 2). In particular, we are interested in identifying the genes that are jointly associated with substance abuse disorder and AD by analyzing two independently collected data sets from raw sequence data.

iii) Despite the proliferation of GWAS tools, detecting epistasis is still challenging. One main limitation of the existing tools is that they can only model linear association signals in the GWAS data. To overcome this challenge, we leverage the power of the deep neural networks and developed a tool, namely Deep Mixed Model (DMM), to model arbitrary interactions of the data.

iv) Cluster analysis has been employed to detect subtypes of complex diseases which is a key task for precision medicine. However, clustering patients based on different

sources/types of data (or called multiview data, e.g., clinical, gene expression, and proteomic data) can be challenging because different data has its own statistical property that is different from other data. Existing approaches that aim to address this problem can yield unfavorable results that largely depend on certain types of data when noise or redundant variables present in the multiview data. We developed a robust multiple kernel k-means clustering approach, called MML-MKKC, and showed that our method can robustly identify true clusters when noise or redundant variables are present in multiview data.

v) Better understanding how corticosteroids (CSs) use affects asthma patients is important for precision treatment of these patients. For this purpose, we analyzed a rigorously characterized adult asthma cohort from the Severe Asthma Research Program (SARP), and developed a multiview strategy which allows us to identify clusters of the asthma subjects with differential response patterns to CS using MML-MKKC. Using this strategy, we identified four clusters of patients showing differential response patterns among the asthma patients; our clusters were validated using an independent SARP test set.

These machine learning tools will provide state-of-the-art and powerful resources for DA research.

(b) Continued existing collaborations and newly initiated ones with NIDA/NIH-funded research projects (FRPs) on genome-wide screening of expression traits, and molecular and clinical markers for drug abuse (DA), other neurological diseases (e.g., AD), and diseases related to DA (e.g., asthma). As we proposed in **Aims 1–3**, we have applied our newly developed machine learning models and algorithms technologies to support the FRPs. These projects are listed below.

Collaborations:

1) We have continued to support the research of the PI **Dr. Michael Vanyukov** (who is a consultant and the PI of the neurogenetics module of FRP1 on Core C (CG4DA) and is also the Scientific Director of the Center for Education and Drug Abuse Research (CEDAR) at U Pitt, supported by FRP1) and help identify genome-wide genetic targets for patients with substance abuse disorders. **Dr. Vanyukov** *is the PI of one of the* ***FRPs*** *listed in our funded P30 Center grant application*.

2) We have continued to support the research of the PI **Dr. Oscar Lopez** (Professor of Neurology and Psychiatry, Director of Alzheimer's Disease Research Center, University of Pittsburgh) and help identify genome-wide genetic targets and molecular markers for patients with Alzheimer's disease (AD). **Dr. Lopez** *is the PI of one of the* ***FRPs*** *listed in our funded P30 Center grant application.*

3) We have continued to support the research of the PI **Dr. Sally Wenzel** (Professor of Medicine and Immunology, Pulmonary Medicine, Department Chair of Environmental and Occupational Health, Director of University of Pittsburgh Asthma Institute at UPMC/UPSOM, and UPMC Chair of Translational Airway Biology) and help identify response patterns to CSs among asthma patients. **Dr. Wenzel** *is also the PI of one of the* ***FRPs*** *listed in our funded P30 Center grant application.*

(c) RESEARCH HIGHLIGHTS & SIGNIFICANT RESULTS (major findings, developments or conclusions)

Core Technology Development (*Aims 1–3*):

1. **Discovering Weaker Genetic Associations Guided by Known Associations (*Wang, et al, BMC Medical Genetics 2019*).**

   The vast amount of genomic data has increased the possibility of discovering new genome-phenotype associations with statistical methods. Following the lead of traditional statistical methods like hypothesis testing, many advanced machine learning methods have been proposed which intent to increase the statistical power by incorporating a variety sources of prior knowledge to the models. However, these methods barely consider the prior knowledge of validated associations, which results in the following limitations of the current association studies: 1) A majority of the newly discovered genetic variants associated with a disease are known variants; and 2) as the validated/known variants have larger effect sizes than other variants with weaker association, the former are easier to discover; as such, much effort has been wasted on discovering the same (known) variants.

   We propose a computational approach CS-LMM that uses the knowledge of validated associations to increase the power of discovery of genetic variants (*Wang, et al, BMC Medical Genetics 2019*). Therefore, statistically, since the searching is conditioned on validated associations with large effect sizes, the weaker signals will be easier to be uncovered. Moreover, the discovered genetic variants will not overlap largely on what is already known. Our simulation experiments show that CS-LMM outperforms other methods in terms of discovering genetic variants with smaller effect sizes. We applied our method to an alcoholism SNP dataset provided by our collaborator Dr. Michael Vanyukov as well as an Alzheimer SNP dataset, and identified a dozen promising genetic variants potentially associating with drug abuse disorders as well as Alzheimer's.

2. **Coupled Mixed Model for Joint Genetic Analysis of Complex Disorders with Two Independently Collected Data Sets.**

   Genome-wide Association studies (GWASs) have contributed to decoding the human genome by uncovering many genetic variations associated with various diseases. Many follow-up investigations involve joint analysis of multiple independently generated GWAS data sets. While most of the computational approaches developed for joint analysis are based on summary statistics, the joint analysis based on individual-level data with consideration of confounding factors remains to be a challenge. We developed a method, called Coupled Mixed Model (CMM), that enables a joint GWAS analysis on two independently collected sets of GWAS data with different phenotypes. The CMM method does not require the data sets to have the same phenotypes as it aims to infer the unknown phenotypes using a set of multivariate sparse mixed models. Moreover, CMM addresses the confounding variables due to population stratification, family structures, and cryptic relatedness, as well as those arising during data collection such as batch effects that frequently appear in joint genetic studies. We evaluate the performance of CMM using simulation experiments, and our method consistently outperforms the competing methods, as shown in Figure 1 (Aims 1 & 2).

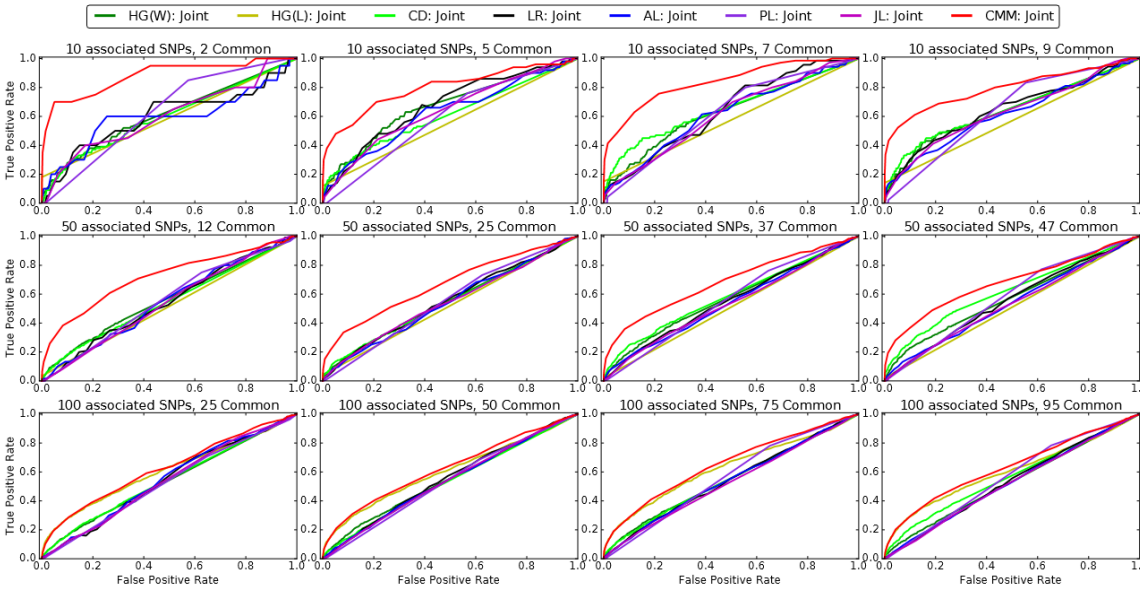


Figure 1. Empirical performance of CMM in comparison to other popular methods in simulation experiments.

3. Deep Mixed Model for Marginal Epistasis Detection and Population Stratification Correction in Genome-Wide Association Studies (*Wang, et al, BMC Bioinformatics 2019*).

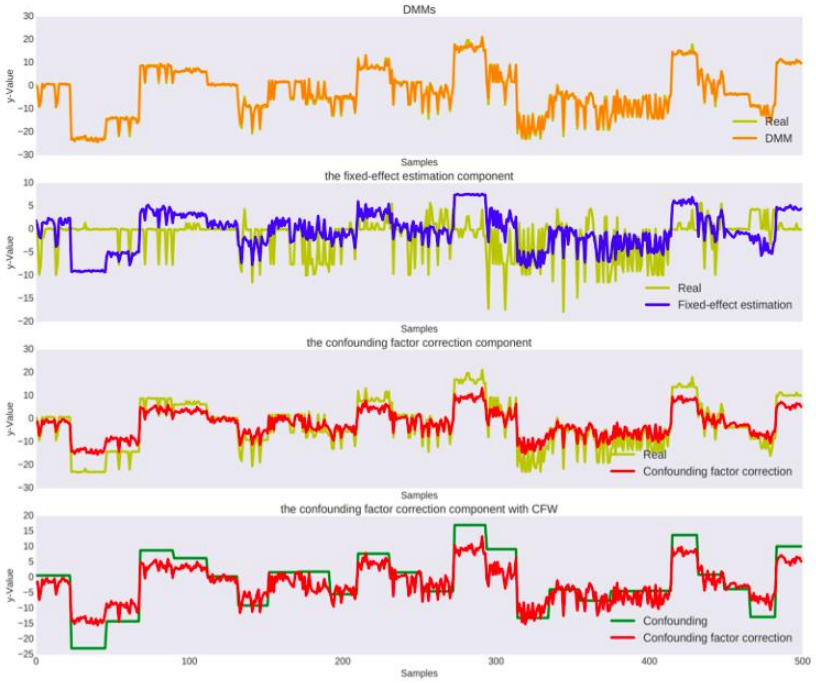


Figure 2. Investigation of the internal working mechanism of DMM.

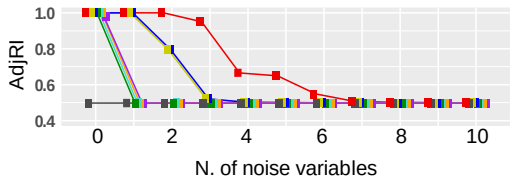
Following the previous study of detecting marginal epistasis signals, and motivated by the universal approximation power of deep learning, we propose a neural network method that can potentially model arbitrary interactions between SNPs in genetic association studies as an extension to the mixed models in correcting confounding factors. Our method, namely Deep

Mixed Model, consists of two components: 1) a confounding factor correction component, which is a large-kernel convolution neural network that focuses on calibrating the residual phenotypes by removing factors such as population stratification, and 2) a fixed-effect estimation component, which mainly consists of an Long-short Term Memory (LSTM) model that estimates the association effect size of SNPs with the residual phenotype. With simulations, we demonstrate the superior performance over the existing methods. We also investigate the internal working mechanism of our proposed DMM (as shown in Figure 2). Our investigation suggests that DMM does not only achieve good end-performance, but also behave as we expected in intermediate steps.

4. **Robust Multiple Kernel k-means Clustering using Min-Max Optimization (*Wu, et al, AJRCCM 2019*).**

Integrating diverse modalities is challenging because data from different sources (also called views) have different statistical properties. To address this problem, multiple kernel learning uses view-specific kernels to capture diverse patterns of multiple views. While supervised multiple kernel learning has been extensively studied, only a few unsupervised approaches have been proposed until recently, among which, multiple kernel k-means clustering is one of

(a)



(b)

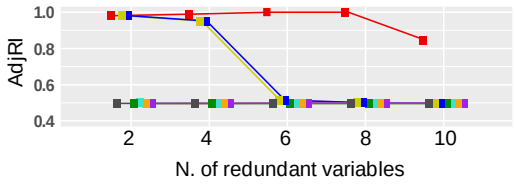


Figure 3: Clustering performance. Adjusted Rand Index (AdjRI) versus (a) the number of noise variables and (b) the number of redundant variables added to one of the views are plotted. The identified clusters are compared to the true clusters.

the commonly used approaches. Existing works employ a minH-minθ (or maxH-maxθ) framework, in which they first find a combination of views that reveals small within-cluster variance, and then find clusters by minimizing such variance. However, noise or redundant variables can make the existing multiple kernel clustering approaches yield unfavorable clusters. In particular, the minH-minθ framework makes these methods ignore the views with compromised signals and find clusters that are largely determined by other views.

To address this problem, we propose a multiple kernel clustering method with the minH-maxθ framework (MML-MKKC) that aims to be robust to the noise and redundant variables. Our simulation results show that our method is more robust to noise (Figure 3a) or redundant (Figure 3b) features than other compared methods. Our method outperforms the compared existing multiple kernel clustering methods and yields clusters by making good use of all views, including the view with the noise or redundant variables.

5. Multiview Clustering Analysis for Identifying Variable Corticosteroid Response Phenotypes in Severe Asthma (*Wu, et al, AJRCCM 2019*).

Corticosteroids (CSs) are the most effective asthma therapy, but responses are heterogeneous and systemic CSs lead to long-term side effects. Therefore, an improved understanding of the contributing factors in CS responses could enhance precision management. Although several factors have been associated with CS responsiveness, no integrated/cluster approach has yet been undertaken to identify differential CS responses.

In order to help understand differential CS responsiveness among subjects with asthma, we developed a multiview strategy (Figure 4) which allows us to identify clusters of asthma subjects with differential response patterns to CS by i) incorporating different types of variables, including both baseline and change variables, into the cluster analysis; and ii) assigning variables into different views based on their clinical importance. We applied our method MML-MKKC using the three-view strategy to a rigorously characterized adult asthma cohort from the National Institutes Heart Lung and Blood Institutes' Severe Asthma Research Program (SARP). The patients in this cohort were studied before and after a standardized systemic CS treatment to characterize their responses.

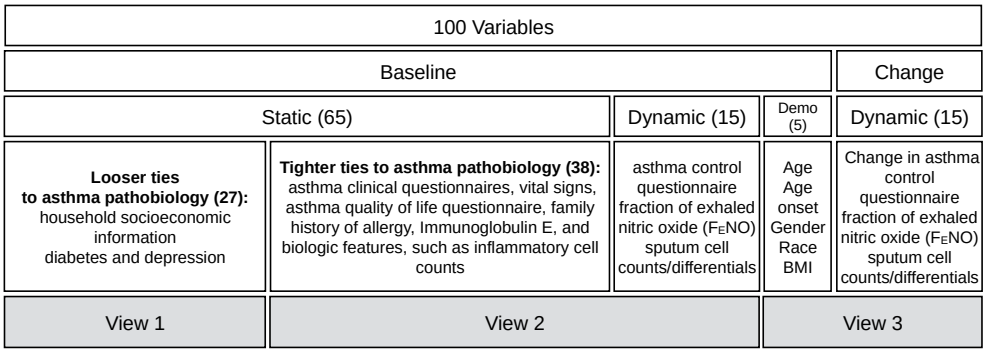


Figure 4. Three-view strategy for identifying variable CS response phenotypes in SARP. The 100 clinical, physiological, inflammatory, and demographic variables from 346 adult participants with asthma in SARP with paired (before and after CS use) sputum data are divided into three views.

We identified four asthma clusters with differential CS responses among 346 asthma patients. The newly identified clusters are validated/replicated using an independent SARP test set. These findings give insight into clinical, biologic and physiologic determinants of CS response patterns that could be mechanistically utilized to better link molecular to clinical responses.

Collaborations with FRPs/Pilot (*Aims 1-3*):

(1) Genome-wide association analysis of alcoholism data (collaboration with Dr. Michael Vanyukov)

To help support Dr. Michael Vanyukov's research, we applied our method CS-LMM to a alcoholism dataset collected from the Center for Education and Drug Abuse Research (CEDAR) at U Pitt. There are 519,138 SNPs collected from 383 subjects with and without drug abuse disorders in this dataset.

Key outcomes/Achievements:

SNPs associated with alcoholism (*Wang, et al, BMC Medical Genetics 2019*). Using CS-LMM, we identified multiple SNPs associated with drug abuse disorders, of which, the top 20 SNPs associated with alcoholism are shown in Table 1. It can be seen that many of them are in the region of gene ALDH3A2, which is a known gene involved in alcohol metabolism and associated with smoking cessation behavior. ENOX1 is also known to be associated with nicotine dependence.

SNP	gene	Chromosome	SNP	gene	chromosome
rs1789891	**ADH1B**	**4**	rs12482570	KCNJ6	21
rs7590720	**PECR**	**2**	rs857975	KCNJ6	21
rs2835872	KCNJ6	21	rs4147544	ADH6	4
rs4478858	SERINC2	1	rs702860	KCNJ6	21
rs1789924	ADH1C	4	rs2835853	KCNJ6	21
rs698	ADH1C	4	rs717859	KCNJ6	21
rs2851300		4	rs11499823	ADH1C	4
rs10483038	KCNJ6	21	rs2835910	KCNJ6	21
rs1344694	PECR	2	rs4355398		4
rs4147536	ADH1B	4	rs2835831	ADH6	4

Table 1: The top 20 SNPs discovered by CS-LMM. The top four SNPs (shown in bold) are the ones that are built into the model to help discover weaker signals. The rest 17 SNPs are novel SNPs associated with alcoholism we discovered using CS-LMM.

(2) Genome-wide association analysis of late-onset Alzheimer's disease data (collaboration with Dr. Oscar Lopez)

To help support Dr. Lopez's research, we applied our method CS-LMM to a late-onset Alzheimer's disease (AD) dataset from Harvard Brain Tissue Resource Center and Merck Research Laboratories in an attempt to detect causal SNPs associated with AD. This dataset contains 555,091 SNPs obtained from 270 AD cases and 270 controls (non-demented subjects).

Key outcomes/Achievements:

We identified SNPs associated with Alzheimer's disease, some of which suggest a potential link between AD and drug abuse disorders (*Wang, et al, BMC Medical Genetics 2019*). Using CS-LMM, we identified multiple SNPs associated with AD, of which, the top 15 novel SNPs associated with AD are shown in Table 2. Interestingly, our literature survey suggests that several SNPs in Table 2 are previously known associated with alcoholism.

SNP	gene	chromosome	SNP	gene	chromosome
rs2075650	**APOE**	**19**	rs12131508	SLC35F3	1
rs157580	**TOMM40**	**19**	rs12506821		4
rs10027926	RGS12	4	rs11485175		1
rs12641989	RGS12	4	rs584507	PRKCQ	10
rs3088231	RGS12	4	rs12563692	ESRRG	1
rs10512523	ABCA9	17	rs6446731		4
rs4076949	SLC35F3	1	rs7984051		13
rs874418	HGFAC	4	rs2327771	ISM1	20
rs6842419	DOK7	4	rs7548651	SLC35F3	1
rs16844383	HGFAC	4	rs4330674	WISP1	8

Table 2: The top 20 SNPs associated with AD discovered by CS-LMM. The top two SNPs (shown in bold) are the ones that are built into the model to help discover weaker signals. The rest 18 SNPs are novel SNPs associated with AD

(3) Genome-wide association analysis of both drug abuse disorder and Alzheimer's disease. (collaboration with Drs. Michael Vanyukov and Oscar Lopez)

To help support Drs. Michael Vanyukov's and Lopez's research, we applied our method CMM for a joint study of substance abuse disorder (SUD) and Alzheimer's disease. Our results show that our method the *CMM* can identify several interesting markers that are jointly associated with both diseases.

Key outcomes/Achievements:

We identify five SNPs that are jointly associated with SUD and Alzheimer's disease.

SNP	SUD rank	AD rank	Chr.	Chr. Position	Gene
rs2131691	1	1	11	26574855	*ANO3*
rs1709317	5	8	2	23536638	*KLHL29*
rs4713797	6	10	6	34490756	*PASCIN1*
rs224534	12	3	17	3583408	*TRPV1*
rs1057744	16	11	14	105150705	*JAG2*

Table 3. SNPs identified to be jointly associated with SUD and Alzheimer's disease.

Table 3 shows the SNPs that our CMM identified to be jointly associated with both SUD and Alzheimer's disease. Interestingly, one of the SNPs reside in *TRPV1*, which is a gene that is predicted independently by the Core B members to be related to drug abuse.

(4) Multiview clustering analysis of Asthma treatment responsiveness. (collaboration with Dr. Sally Wenzel)

To help support Dr. Wenzel's research, we applied our method MML-MKKC to an asthma clinical dataset collected from the participants in Severe Asthma Research Program (SARP) (provided by Dr. Wenzel). There are 346 asthma patients and 100 clinical, physiologic, inflammatory, and demographic variables in this dataset. Our results show that our method *MML-MKKC* identified four response patterns of asthma patients to corticosteroids (CSs).

Key outcomes/Achievements:

We identified four response patterns of asthma patients to CSs (*Wu, et al, AJRCCM 2019*). Figure 5 shows the four asthma clusters with different CS responses among 346 asthma patients. Clusters 1 and 2 consisted of young, modestly CS-responsive individuals with allergic asthma and relatively normal lung function, separated by contrasting sputum neutrophil and macrophage percentages after CS treatment. The subjects in cluster 3 had late-onset asthma and low lung function, high baseline eosinophilia, and the greatest CS responsiveness. Cluster 4 consisted primarily of young, obese females with severe airflow limitation, little eosinophilic inflammation, and the least CS responsiveness. The top 12 baseline variables were identified, and the clusters were validated using an independent SARP test set. These findings give insight into clinical, biologic and physiologic determinants of CS response patterns that could be mechanistically utilized to better link molecular to clinical responses.

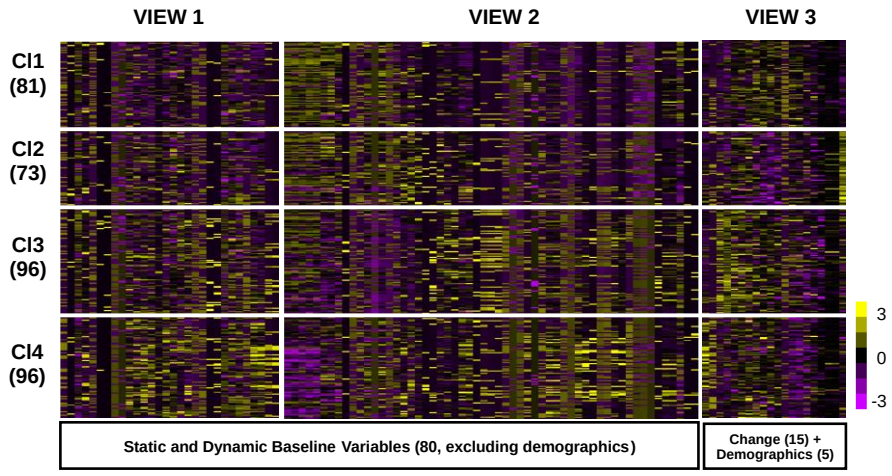


Figure 5: Heatmap of the four clusters identified among the 346 patients by the MML-MKKC method. Rows represent the patients and columns represent variables in each view.

RESULTS DISSEMINATED TO COMMUNITIES OF INTEREST

(1) Presentations and posters were presented in annual society meetings

1. Wang H, Wu Z, Xing EP. Removing Confounding Factors Associated Weights in Deep Neural Networks Improves the Prediction Accuracy for Healthcare Applications. Oral Presentation in Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 2019.

2. Wang H, Liu X, Tao Y, Ye W, Jin Q, Cohen WW, Xing EP. Automatic Human-like Mining and Constructing Reliable Genetic Association Database with Deep Reinforcement Learning. Poster presentation in Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 2019.

(2) Invited Talks

1. Broad Institute Next Generation in Biomedicine Symposium (Boston, MA, USA) Sept. 9, 2019

   Talk title: *"Dealing with Confounding Factors in Deep Learning"* presented by **Haohan Wang** (Carnegie Mellon University, USA)

2. Department of Biomedical Informatics Colloquium Series (Pittsburgh, PA, USA) Sept. 6, 2019

   Talk title: *"Deep Learning over Heterogeneous Data: a challenge, a solution, and an application to Poly(A) signal prediction"* presented by **Haohan Wang** (Carnegie Mellon University, USA)

PRODUCTS

Publications

12. Wu X, Xie S, Wang L, Fan P, Ge S, Xie XQ, Wu W*. A computational strategy for finding novel targets and therapeutic compounds for opioid dependence. PLoS One. 2018 Nov 7;13(11):e0207027. doi: 10.1371/journal.pone.0207027. eCollection 2018. PMID:30403753. PMCID: PMC6221321.
13. Bertholomey ML, Stone K, Lam TT, Bang S, Wu W, Nairn AC, Taylor JR, Torregrossa MM*. Phosphoproteomic analysis of the amygdala response to adolescent glucocorticoid exposure reveals G-protein coupled receptor kinase 2 as a target for reducing motivation for alcohol. Proteomes. 2018 Oct 12;6(4). pii: E41. doi: 10.3390/proteomes6040041. PMID:30322021. PMCID: PMC6313880.
14. Wang H, Lengerich BJ, Aragam B, Xing EP. Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. Bioinformatics. 2018 Sep 1;35(7):1181-7.
15. Wang H, Aragam B, Xing EP. Variable selection in heterogeneous datasets: A truncated-rank sparse linear mixed model with applications to genome-wide association studies. Methods. 2018 Aug 1;145:2-9.
16. Wang H, Liu X, Xiao Y, Xu M, Xing EP. Multiplex confounding factor correction for genomic association mapping with squared sparse linear mixed model. Methods. 2018.
17. Marchetti-Bowick M, Yu Y, Wu W, Xing EP*. A penalized regression model for the joint estimation of eQTL associations and gene network structure. The Annals of Applied Statistics. 2019;13(1):248-70.
18. Wang H, Wu Z, Xing EP. Removing Confounding Factors Associated Weights in Deep Neural Networks Improves the Prediction Accuracy for Healthcare Applications. In Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 2019 (Vol. 24, p. 54). NIH Public Access.
19. Wang H, Liu X, Tao Y, Ye W, Jin Q, Cohen WW, Xing EP. Automatic Human-like Mining and Constructing Reliable Genetic Association Database with Deep Reinforcement Learning. InPacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 2019 (Vol. 24, pp. 112-123). NIH Public Access.
20. Wang H, Yue T, Yang Y, Wu W, Xing EP. Deep Mixed Model for Marginal Epistasis Detection and Population Stratification Correction in Genome-Wide Association Studies, BMC Bioinformatics. Accepted.
21. Wang H. Vanyukov MM, Xing EP, Wu W. Discovering Weaker Genetic Associations Guided by Known Associations, BMC Medical Genetics. Accepted.
22. Wang H. Lu C. Wu W. Xing EP, Graph-structured Sparse Mixed Models for Genetic Association with Confounding Factors Correction, BIBM 2019.
23. Wang H, Pei F, Vanyukov MM, Bahar I, Wu W, Xing EP. Coupled Mixed Model for joint genetic analysis of complex disorders from independently collected data sets: application to Alzheimer's disease and substance use disorder. Submitted to Bioinformatics. (https://www.biorxiv.org/content/10.1101/336727v2.article-metrics)
24. Wu W, Bang S, Bleecker ER, Castro M, Denlinger L, Erzurum SC, Fahy JV, Fitzpatrick AM, Gaston BM, Hastie AT, Israel E, Jarjour NN, Levy BD, Mauger DT, Meyers DA, Moore WC, Peters M, Phillips BR, Phipatanakul W, Sorkness RL, Wenzel SE. Multiview Cluster Analysis Identifies Variable Corticosteroid Response Phenotypes in Severe Asthma. American Journal of Respiratory and Critical Care Medicine 2019 Jun 1;199(11):1358-1367.
25. Bang S, Yu Y, Wu W. Robust Multiple Kernel k-means Clustering using Min-Max Optimization. Submitted to the AAAI conference.

Technologies or Techniques

GenAMap: An Visual Analytics Software Platform for eQTL and GWAS Analysis.
https://github.com/blengerich/GenAMap

Precision Lasso: Accounting for Correlations and Linear Dependencies in High-Dimensional Genomic Data.
https://github.com/HaohanWang/thePrecisionLasso

CS-LMM: Discovering Weaker Genetic Associations with Validated Association.
https://github.com/HaohanWang/CS-LMM

Deep Mixed Model: Marginal Epistasis Detection and Population Stratification Correction in Genome-Wide Association Studies
https://github.com/HaohanWang/DMM

MKKC: An R-package For Multiple Kernel K-means Clustering.
https://seojinbang.github.io/MKKC/

CORE D

Pilot/Feasibility Projects

EXECUTIVE SUMMARY

Xiang-Qun Xie, PhD,MBA
Ivet Bahar, PhD
Eric Xing PhD

AIMS

Specific Aim 1. To support new P/FPs and identify other promising P/FPs for incorporation into CDAR.

Specific Aim 2. To support the success of P/FPs in their research and their efforts for securing funding, and their promotion to FRP status.

Specific Aim 3. To continue building a network that integrates experimental and computational DA researchers locally and nationwide

ACCOMPLISHMENTS

During the first funding cycle, our P/F program has supported 7 pilot projects on DA research, which has led to 18 publications[1-18] (**Table 1**). The P/F program has recruited two early-stage investigators new to DA research: Dr. Zhiwei Feng, Assistant Professor at Pitt-SOP and Dr. Junmei Wang, Associate Professor at Pitt-SOP. They are now key personnel (**Core A** Co-I and co-PI/Coordinator, respectively) in the proposed new funding cycle of CDAR. Through the P/F program, CDAR also worked closely to help seven junior faculty on nine funding applications (**Table 2**). Among them, Drs. V. Blair Journigan and Zachary Freyberg are PIs of new P/F projects (**P/FPs**) in the new term. The comprehensive support provided by CDAR to early-stage investigators facilitates their long-term research in the DA field. As a successful example, an R01 grant has been awarded to Dr. Torregrossa, the PI of a previous P/FP, to continue her study on cocaine addiction. In addition, the Center was successful recruiting and supporting underrepresented minorities. We have supported Dr. Hernandez (Hispanic, female) for an Alzheimer's Association Fellowship application, Dr. Torregrossa (female) for her R01 grant application, and Dr. Journigan (female) for her K01 grant application (**Table 2**). The Center also has a good representation of female investigators as PI, Co-I, and Coordinator (Bahar, Wu, and Cheng) as well as FRP PIs.

Table 1. Accomplishments of CDAR P/FPs in the First Funding Cycle

Investigator (*Affiliation*)	Project Title	Publication
Peng Yang (*Pitt*)	Designing novel functional chemical probes with high cannabinoid receptor CB1/CB2 selectivity and specificity as probes for studying of cannabinoid related pathways	1-10
Mingfeng Bai (*Pitt*)	Development of CB2 receptor (CB2R)-targeted theranostic PET agent for brain cancer research	11
Yong Wan (*Northwestern*)	Targeting the interplay between KLF4 and PRMT5 in carcinogenesis	12-14
Youhua Liu (*Pitt*)	Targeted inhibition of the type 2 cannabinoid receptor	15
Min Xu (*CMU*)	Development of deep learning-based subdivision approach for large scale macromolecules structure recovery from electron cryo tomograms	16-17
Mary Torregrossa (*Pitt*)	Identification of key proteins in substance addition pathways as drug targets for potential treatment	18
Inmaculada Hernandez (*Pitt*)	Patient and system-level predictors of adherence to oral anticoagulation	-

Table 2. Grants Involving Early-Stage Investigators Supported by the CDAR P/F Program

Investigator (*Affiliation*)	Grant Title (Funding Source)	Status
Mary Torregrossa (*Pitt*)	Mechanisms Regulating Cocaine Memory Strength (NIDA R01DA042029)	Awarded
Inmaculada Hernandez (*Pitt*)	Claims Data Mining to Predict Side Effects of Anti-dementia Drugs (Alzheimer's Association AARGD-17-500234)	Awarded
	Patient, Systems-Level Determinants of Oral Anticoagulation in Atrial Fibrillation (NHLBI K01HL142847)	Awarded
Lirong Wang (*Pitt*)	Chemogenomics Systems Pharmacology Approach for TBI and AD Research (Department of Defense W81XWH-16-1-049)	Awarded
	Synaptic Resilience to Psychosis in Alzheimer Disease (NIME R01MH116046)	Awarded
V. Blair Journigan (*Marshall University*)	Somatosensory-targeting Probes for Neuropathic Pain (NIH K01)	Pending
Zachary Freyberg (*Pitt*)	A Multidisciplinary Approach to Decipher Dopamine D2R Signaling (NIH R35)	Pending
Bing Liu (*Pitt*)	New Therapies for Liver Fibrosis and Hyperproliferation in Alpha1-AT Deficiency (NIDDK P01DK096990)	Awarded
	Signature Directed Sequential Delivery of Small Molecule Radiation Mitigators and Probiotics (NIAID P01)	Pending

P/FPs Selected for the New Funding Cycle

We have already identified five P/FPs for the new funding cycle (**Table 3.**), with priority for those PIs from underrepresented racial and ethnic groups. We provide details on the PIs, their career status, and an overview of their research goals, and how they will interact with and benefit from interactions with CDAR investigators.

Table 3. P/FPs Selected for the New Funding Cycle.

P/FP	Investigator (*Affiliation*)	Grant Title	Cores
1	Ying Xue (*Pitt*)	Substance use disorder risk stratification by machine learning algorithms and Bayesian causal network models	**A**, C
2	V. Blair Journigan (*Marshall U*)	Novel ligands for TRPM8 menthol receptor for smoking cessation	**A**, B
3	Zachary Freyberg (*Pitt*)	Direct visualization of morphological and structural alterations in dopamine neurons caused by drugs of abuse	**B**, C
4	Min Xu (*CMU*)	Developing deep learning approaches for analyzing Cryo-EM imaging data for drug abuse research	**C**, B
5	Scott Malec (*Pitt*)	Literature-based discovery informing graphical causal modeling for drug repurposing	**A**

We here use two examples of F/FPs to demonstrate how the P/FPs can benefit from CDAR center

P/FP1: "Substance use disorder risk stratification by machine learning (ML) algorithms and Bayesian causal network models" (Ying Xue, PhD, Assistant Professor, School of Pharmacy, Pitt)

Specific Aims. Dr. Xue is a new investigator at Pitt, School of Pharmacy, with research focus on substance use disorder (SUD) risk stratification. The objective of this proposal is to address the role of individual clinical characteristics and other factors on SUD. Specifically, Dr. Xue will apply Bayesian models to identify key relationships and combination of factors, which can be further used to assess the probabilistic risk of SUD given demographic factors and comorbid conditions. She will also quantify to what extent the factors' variation observed in individuals is attributable to SUD. The project will benefit from the support of Cores A and C and will focus on evaluating time-invariant and time-varying factors including demographics, clinical data, region of residence and psychotic experiences that relate to the development of SUD over time. By doing so, the P/FP PI and **Core A** in conjunction with **Core C** will leverage the richness of causal networks to identify causal relationships between different factors and SUDs. Building on the results from Bayesian causal networks, the P/FP PI will further develop an algorithm to identify patients with high risk of SUD. To achieve this goal, two specific aims are proposed below:

Aim 1: To characterize the causal relationships between time-invariant and time-varying patients- and system-level factors among SUD patients.

Aim 2: To develop ML algorithms for estimating the individual risk of SUD as well as expected future healthcare utilization.

Background and Significance. SUD remains an increasing global public health concern, resulting in substantial socioeconomic burden. A recent report from the NIDA states that the annual cost from substance abuse in the United States, related to healthcare, crime, and reduced productivity exceeds $700 billion. More than that, SUD has been listed among the top ten non-genetic causes of death (source: https://www.drugabuse.gov). Although, significant academic and health risks associated with SUD have been well recognized, the development of systematic approaches for disease management in population with SUD has been hampered by the limited knowledge of disease/disorder mechanics compared to other diseases/conditions such as cancer and heart disease. As a result, identifying which factors are most robustly linked to SUD is critical for SUD prevention and effective interventions design.

The ability to develop effective models for SUD management is limited by substantial evidence gaps that this pilot project proposes to addresses: (1) While there is extensive literature reporting the comorbidities and impact on SUD, it should be noticed that most of the work has been done using regression modeling and clustering, and these methods suffer from limitations with respect to their ability to codify complex nonlinear relationships, ingest and model large sample sizes, and provide transparent outputs to users. (2) Prediction of substance misuse patterns is still lacking owing to the large number and inter-correlation of predictors along with their changing salience during development (time-varying factors), in other words, the strength of those association may change over time. Hence, we will develop more powerful analytic methods that have the potential to address the current shortcomings in exploring the inherent complexity of comorbidities in SUD population toward developing an individualized model of risk stratification.

Preliminary Work. In our previous research, we used ML algorithms to determine the accuracy of forecasting SUD based on psychological characteristics of children and adolescents (**Figure 1**). Longitudinal data (N = 700) from the Center for Education and Drug Abuse Research was analyzed at baseline and subsequent follow-ups for predicting SUD. From a pool of approximately 1000 questionnaire items, ML feature selection techniques, which are data-driven and free of any assumption or investigator bias, were adopted to select 30 psychological questionnaire items which best predicted SUD. These items were further used to predict the particular substance use severity (SUS) trajectory during adolescence and early adulthood. The ROC AUCs of the ML models for predicting SUD at different ages are 0.74 at ages 10-12 and 12-14, and progressively increased to 0.78, 0.83, and 0.86 at the respective ages 16, 19, and 22 (**Figure 1**). ML methodology is heuristic for delineating and scaling the vulnerability characteristics associated with SUD liability as well as developing liability measurement tools to inform targeted prevention.

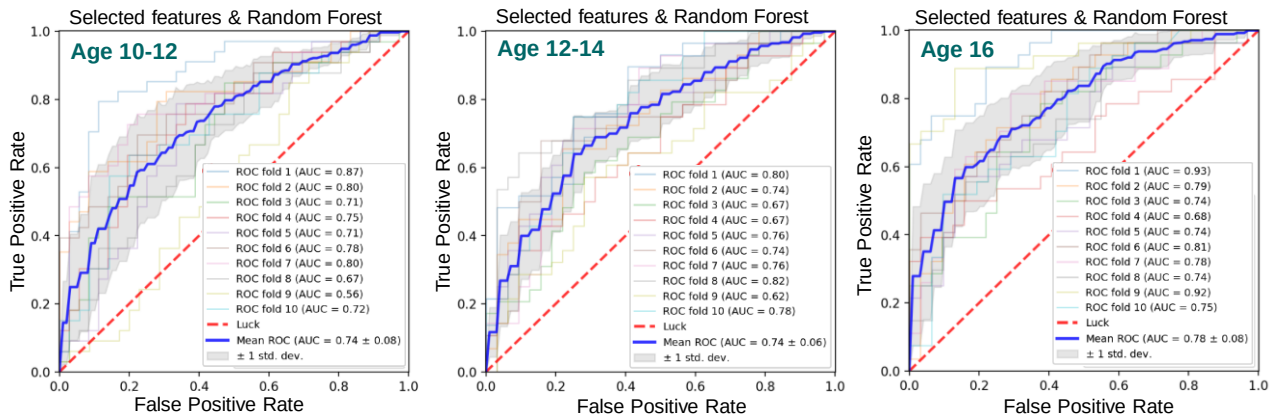


Figure 1 Random forest with feature selection predicted the SUB liability among children and adolescents.

Computations and Benefits from the CDAR Center. The proposed P/F research will benefit from Cores A and C through the following collaboration: (i) to identify what patient characteristics are causally associated with SUD overtime, the P/FP PI will build on the experience of Core C in building Bayesian networks. These advanced models will enable us to develop Markov chains capturing the causal relationship between patient time-invariant characteristics, such as race or gender, system-level characteristics including region of residence, time-varying factors such as psychotic experiences (auditory and visual hallucinations, and five delusions) overtime. (ii) Building on the results from these models and on the experience of trajectory analysis in Core A, the P/FP PI will develop an algorithm to predict patients at highest risk of SUD.

Experimental Validations by P/FP Investigator. The proposed computational studies will be experimentally validated by the P/FP PI. In her research plan, she will train a series of Bayesian network models to estimate individual risk of SUD as well as expected future healthcare utilization. The collaboration with the CDAR team members has the potential for developing more advanced tools for SUD prevention, prediction and management, which can lead to future grants to establish the P/FP PI as an independent scientist. The P/FP PI will also spend significant time focusing on SUD research during the training period learning important computational tools.

P/FP2: "Novel ligands for TRPM8 menthol receptor for smoking cessation" (V. Blair Journigan, PhD, Assistant Professor of Medicinal Chemistry, Department of Pharmaceutical Sciences, Marshall University)

Specific Aims. Dr. Journigan is a new investigator at Marshall University in the Department of Pharmaceutical Sciences. The Journigan lab is focused on the discovery of novel small molecule chemical probes for the transient receptor potential melastatin 8 (TRPM8) ion channel, commonly known as the menthol receptor, to uncover novel therapies for nicotine addiction and neuropathic pain. The project will benefit from the expertise and resources developed in Cores A and B**.** The P/FP PI, assisted by Core A-B members, will first focus on modeling stimuli-dependent structural changes resulting in activation of the channel, to gain a structure-based understanding of conformational rearrangements that may play a role in opening/closing. Subsequent work will focus on docking and molecular dynamics (MD) simulations of TRPM8 agonist and antagonist ligands in our human TRPM8 homology model based on the avian cryo-EM structure (PDB 6BPQ), including our novel antagonist scaffolds, to gain a structure-based understanding of the molecular determinants for ligand recognition. The Journigan lab is interested in elucidating a pharmacophore model for TRPM8 ligands, based on their predicted binding epitopes in the putative orthosteric site, and the tool *Pharmmaker* newly developed by Core B as part of the highly versatile and broadly used *ProDy* API, will be utilized. These computational studies will enable structure-based design of novel menthol-based chemotypes. Targets identified in silico will be synthesized by the P/FP PI, and evaluated in binding, calcium flux and whole-cell patch clamp assays in cross-disciplinary collaborations with her laboratory. The aims of this project are:

Aim 1. To model temperature-, ligand-, and voltage-dependent structural changes responsible for opening and closing of the TRPM8 channel.

Aim 2. To determine the molecular basis for small molecule ligand recognition for agonist and antagonist profiles, guided by molecular docking and MD simulations in our hTRPM8 homology model.

Aim 3. To design and synthesize high affinity, selective menthol-based TRPM8 chemical probes, via molecular docking and MD simulations, iterative SAR studies, and TargetHunter, DrugGui and *Pharmmaker* technologies.

Background and Significance. The monoterpenoid (α)-menthol, a non-selective agonist of transient receptor potential melastatin 8 (TRPM8), is a common and popular additive in cigarettes, and found in

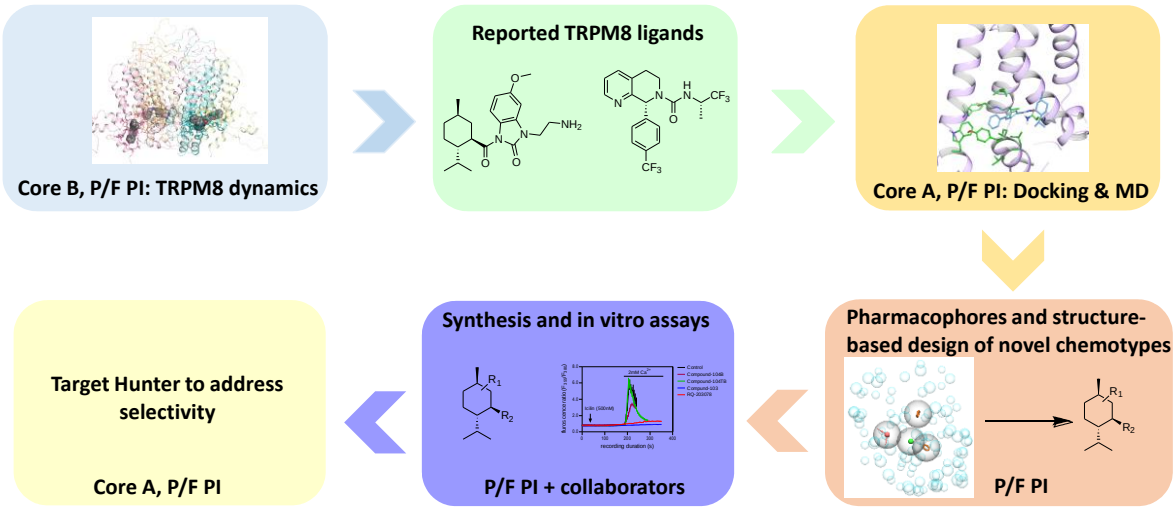


Figure 2. Computational approaches and workflow to identify novel TRPM8 ligands

90% of tobacco products regardless of being labeled either menthol or non-menthol. Recent FDA evaluations of the health risks of menthol cigarettes state that menthol has drug-like characteristics associated with increased addiction and dependence to nicotine. Clinical and epidemiological studies of menthol vs. non-menthol smokers have found that menthol smokers have consistently lower abstinence rates and increased dependence to cigarette smoking, suggesting that menthol in cigarettes supports facilitation and maintenance of smoking behaviors. In rodents, menthol is a counter-irritant against the harsh effect of cigarette smoke and oral nicotine via its TRPM8 agonist activity, likely allowing for longer smoke retention and giving greater exposure to nicotine. TRPM8 activators such as oral menthol, cold temperature (11^{o}C) and partial agonist WS-23 increase nicotine intravenous self-administration, while oral menthol induces a nicotine extinction burst, and reinstates extinguished nicotine-seeking behavior via its cooling sensation in rats, suggesting an interplay between TRPM8 and nAChRs beyond the sensory effects discussed above. We **hypothesize** that menthol-based chemical probes, specifically antagonists, may be useful tools to investigate TRPM8-mediated effects on nicotine pharmacology, and inhibit the counterirritant and reinforcing effects of menthol. In turn, these chemical probes may uncover a novel target for nicotine addiction and provide small molecule lead scaffolds for smoking cessation medications. Currently, no high affinity (in the nanomolar range), selective TRPM8 ligands based on the natural ligand menthol exist to investigate these effects.

Computations and Benefits from the CDAR Center. The proposed P/F research will benefit from Cores A and B of the CDAR Center through the following collaborations: (i) To determine conformational rearrangements resulting from cold temperature, ligand, and voltage stimuli leading to TRPM8 channel activation. The P/FP PI and Core A propose to work together to explore the dynamics of channel opening and closing, which will allow for a mechanistic understanding of activation/inactivation and ion conductance. (ii) To perform docking and MD simulations on reported TRPM8 scaffolds, followed by structure-based pharmacophore generation. This work will be done by the P/FP PI and Cores A and B, using Amber, Desmond, the Schrödinger Small Molecule Drug Discovery suite as well as DrugGui and *Pharmmaker*. An understanding of ligand recognition in our human TRPM8 homology model will allow for the rational design of novel TRPM8 chemotypes. (iii) To perform docking on designed TRPM8 ligands, in order to prioritize their synthesis, and screen identified hit compounds for off-target activity. This work will be done by the P/FP PI and Cores A and B, using Amber, Desmond, TargetHunter, and BalestraWeb. Specifically, designed ligands will be prioritized for synthesis by the P/FP PI based on their docking scores within our model, then evaluated in vitro for binding and functional activity in her collaborators' laboratories. SAR analysis will be performed in an iterative manner. Compounds displaying binding and functional activity less than 1 μM will be profiled for potential off-target activity. The Core A PI (Dr. Xie) has expertise in computational drug discovery for the TRP channel TRPV1 (capsaicin receptor), including the molecular docking studies proposed. The P/FP PI will also use her expertise in structure-based drug discovery and design.

Experimental Validations by P/FP Investigator. The proposed computational studies will be experimentally validated by the P/FP PI. She will use a synthetic medicinal chemistry approach coupled with detailed pharmacological characterization in human TRPM8 cell lines, including direct *in vitro* binding studies by NMR, calcium flux assays, and whole-cell patch clamp electrophysiology. These studies will be useful to identify high affinity, selective menthol-based TRPM8 chemical probes. She proposes to use established synthetic methodology in the menthol literature to access the targets of interest and for analog synthesis. The collaboration with CDAR team has the potential for identifying small molecule menthol-based probes that can be used as tools to study the role of menthol in facilitating nicotine addiction, that can lead to future grants as an independent scientist. The proposed experiments could also lead to novel smoking cessation medications, by virtue of their small molecule templates. The P/FP PI will also spend significant time during the training period learning important computational tools for her research.

PLANS FOR NEXT REPORTING PERIOD

The Core D leaders will continue to provide supports to facilitate and accelerate the research activities of new and burgeoning investigators in the field of Drug Abuse Research (DAR) by helping them incorporate cutting-edge computational tools into their research programs. Core D will work closely with P/F Project

PIs to help them advance their research and secure funding for their work. On a broader level, Core D will facilitate the effective dissemination of CDAR data and tools and promote their use to the broader DAR community, and will also help identify new projects to be included in the P/FP Program.

REFERENCES

1. Alqarni, M., Myint, K.Z., Tong, Q., Yang, P., Bartlow, P., Wang, L., Feng, R. & Xie, X.Q. Examining the critical roles of human CB2 receptor residues Valine 3.32 (113) and Leucine 5.41 (192) in ligand recognition and downstream signaling activities. *Biochem Biophys Res Commun* **452**, 334-339 (2014).
2. Feng, Z., Alqarni, M.H., Yang, P., Tong, Q., Chowdhury, A., Wang, L. & Xie, X.Q. Modeling, molecular dynamics simulation, and mutation validation for structure of cannabinoid receptor 2 based on known crystal structures of GPCRs. *J Chem Inf Model* **54**, 2483-2499 (2014).
3. Feng, Z., Pearce, L.V., Xu, X., Yang, X., Yang, P., Blumberg, P.M. & Xie, X.Q. Structural insight into tetrameric hTRPV1 from homology modeling, molecular docking, molecular dynamics simulation, virtual screening, and bioassay validations. *J Chem Inf Model* **55**, 572-588 (2015).
4. Xie, X.Q., Yang, P., Zhang, Y., Zhang, P., Wang, L., Ding, Y., Yang, M., Tong, Q., Cheng, H., Ji, Q., McGuire, T., Yuan, W., Cheng, T. & Gao, Y. Discovery of novel INK4C small-molecule inhibitors to promote human and murine hematopoietic stem cell ex vivo expansion. *Sci Rep* **5**, 18115 (2015).
5. Feng, Z., Pearce, L.V., Zhang, Y., Xing, C., Herold, B.K., Ma, S., Hu, Z., Turcios, N.A., Yang, P., Tong, Q., McCall, A.K., Blumberg, P.M. & Xie, X.Q. Multi-functional diarylurea small molecule inhibitors of TRPV1 with therapeutic potential for neuroinflammation. *AAPS J* **18**, 898-913 (2016).
6. Teramachi, J., Silbermann, R., Yang, P., Zhao, W., Mohammad, K.S., Guo, J., Anderson, J.L., Zhou, D., Feng, R., Myint, K.Z., Maertz, N., Beumer, J.H., Eiseman, J.L., Windle, J.J., Xie, X.Q., Roodman, G.D. & Kurihara, N. Blocking the ZZ domain of sequestosome1/p62 suppresses myeloma growth and osteoclast formation in vitro and induces dramatic bone formation in myeloma-bearing bones in vivo. *Leukemia* **30**, 390-398 (2016).
7. Xie, X.-Q., Wang, L., Wang, J., Xie, Z., Yang, P. & Ouyang, Q. In silico chemogenomics knowledgebase and computational system neuropharmacology approach for cannabinoid drug research. in *Neuropathology of Drug Addictions and Substance Misuse* 183-195 (Elsevier, 2016).
8. Bian, Y., Feng, Z., Yang, P. & Xie, X.Q. Integrated in silico fragment-based drug design: case study with allosteric modulators on metabotropic glutamate receptor 5. *AAPS J* **19**, 1235-1248 (2017).
9. Cha-Molstad, H., Yu, J.E., Feng, Z., Lee, S.H., Kim, J.G., Yang, P., Han, B., Sung, K.W., Yoo, Y.D., Hwang, J., McGuire, T., Shim, S.M., Song, H.D., Ganipisetti, S., Wang, N., Jang, J.M., Lee, M.J., Kim, S.J., Lee, K.H., Hong, J.T., Ciechanover, A., Mook-Jung, I., Kim, K.P., Xie, X.Q., Kwon, Y.T. & Kim, B.Y. p62/SQSTM1/Sequestosome-1 is an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis. *Nat Commun* **8**, 102 (2017).
10. Chen, S., Feng, Z., Wang, Y., Ma, S., Hu, Z., Yang, P., Chai, Y. & Xie, X. Discovery of novel ligands for TNF-alpha and TNF receptor-1 through structure-based virtual screening and biological assay. *J Chem Inf Model* **57**, 1101-1111 (2017).
11. Zhang, S., Jia, N., Shao, P., Tong, Q., Xie, X.Q. & Bai, M. Target-selective phototherapy using a ligand-based photosensitizer for type 2 cannabinoid receptor. *Chem Biol* **21**, 338-344 (2014).
12. Hu, D., Gur, M., Zhou, Z., Gamper, A., Hung, M.C., Fujita, N., Lan, L., Bahar, I. & Wan, Y. Interplay between arginine methylation and ubiquitylation regulates KLF4-mediated genome stability and carcinogenesis. *Nat Commun* **6**, 8419 (2015).
13. Zhou, Z., Luo, A., Shrivastava, I., He, M., Huang, Y., Bahar, I., Liu, Z. & Wan, Y. Regulation of XIAP turnover reveals a role for USP11 in promotion of tumorigenesis. *EBioMedicine* **15**, 48-61 (2017).
14. Zhou, Z., Feng, Z., Hu, D., Yang, P., Gur, M., Bahar, I., Cristofanilli, M., Gradishar, W.J., Xie, X.Q. & Wan, Y. A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine* **44**, 98-111 (2019).
15. Zhou, L., Zhou, S., Yang, P., Tian, Y., Feng, Z., Xie, X.Q. & Liu, Y. Targeted inhibition of the type 2 cannabinoid receptor is a novel approach to reduce renal fibrosis. *Kidney Int* (2018).

16. Xu, M., Chai, X., Muthakana, H., Liang, X., Yang, G., Zeev-Ben-Mordehai, T. & Xing, E.P. Deep learning-based subdivision approach for large scale macromolecules structure recovery from electron cryo tomograms. *Bioinformatics* **33**, i13-i22 (2017).
17. Liu, C., Zeng, X., Lin, R., Liang, X., Freyberg, Z., Xing, E.P. & Xu, M. Deep learning based supervised semantic segmentation of electron cyro-subtomograms. *IEEE International Conferences on Image Processing (ICIP)*, 1578-1582 (2018).
18. Bertholomey, M.L., Stone, K., Lam, T.T., Bang, S., Wu, W., Nairn, A.C., Taylor, J.R. & Torregrossa, M.M. Phosphoproteomic analysis of the amygdala response to adolescent glucocorticoid exposure reveals G-protein coupled receptor kinase 2 as a target for reducing motivation for alcohol. *Proteomes* **6**(2018).