OMB No. 0925-0001 and 0925-0002 (Rev. 03/2020 Approved Through 02/28/2023)

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.  
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Lu, Xinghua

eRA COMMONS USER NAME (credential, e.g., agency login): luxing

POSITION TITLE: Professor, Biomedical Informatics

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

| INSTITUTION AND LOCATION | DEGREE  (if applicable) | Completion Date  MM/YYYY | FIELD OF STUDY |
| --- | --- | --- | --- |
| Shandong Medical University | M.D. | 1984 | Medicine |
| Shandong Medical University | M.S. | 1988 | Cardiology |
| University of Connecticut Health Center | Ph.D. | 1998 | Pharmacology |
| University of Pittsburgh | Certificate | 2003 | Biomedical Informatics |

# A. Personal Statement

I have long-standing interest and experience in studying cellular signaling system throughout a decade of “wet-lab” research as a pharmacologist and two decades of “dry-lab” research as a computational biologist using machine learning and artificial intelligence approaches. I am the Director of the Cancer Pathway Program in the Center of Causal Discovery, an NIH National Center of Excellence in Big Data to Knowledge (BD2K). For last decade, my research concentrates on developing computational methods for studying cancer pathways and cancer pharmacogenomics, particularly using Bayesian causal network discovery and deep learning approaches. Our group has developed the tumor-specific causal inference model that enable to discover the candidate genomic alterations that drive phenotypic changes observed in individual tumors, which lays a foundation for discovering tumor-specific immune evasion mechanisms. We also have extensive experience in developing probabilistic topic models, which can be used to identify gene expression modules from single cell data. With a unique combination of expertise in pharmacology, signal transduction pathway modeling, immunology and machine learning, my group is well prepared to carry out the proposed cross-domain research. Our translational cancer informatics research led to identification of molecular signatures highly predictive of breast cancer outcome, and we won the second place in the [DREAM 7 Challenge](https://sagesynapse.wordpress.com/2012/11/01/breast-cancer-challenge-team-pitttransmed-places-second-for-metabric-phase-of-the-challenge/); our Bayesian network-based model for studying cellular signaling pathway won the 1st place in the SBV Trans-species Network Inference Challenge ( <https://sbvimprover.com/challenge-2/overview> ); our AI methods for clinical decision support won tied 1st place in the [BioBank Disease AI Challenge](https://rc.partners.org/news-events/announcements/winners-announced-first-biobank-disease-challenge). As the contact PIs of this project, I will be responsible for overall project, including participating in design of computational algorithms, data analysis, writing manuscripts and report to NIH.

**Relevant publications:**

1. Lu, S., Lu, KN, Cheng, S., Ma, X., Nystrom, N., **Lu, X**. (2015) Identifying driver genomic alterations in cancers by searching minimum-weight, mutually exclusive sets. ***PLoS Computational Biology*** 11(8):e1004257
2. Chen, L., Cai, C., Chen, V., and **Lu, X** (2016) Learning a hierarchical representation of the yeast transcriptomic machinery using an autoencoder model. ***BMC Bioinformatics*** 17(Suppl 1):S9

[PMC4895523](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4895523/).

1. Lu, S., Cai, C., Yan, G., Zhou, Z., Wan, Y., Chen, V., Chen, L., Cooper, GF., Obeid, LM., Hannun, YA., Lee, AV., **Lu, X.** (2016) Signal-oriented pathway analyses reveal a signaling complex as a synthetic lethal target for p53 mutations. ***Cancer Research.* 76** (23), 6785-6794
2. Chen, V., Paisley, J., and **Lu, X** (2017) Revealing common disease mechanisms shared by tumors of different tissues of origin through semantic representation of genomic alterations and topic modeling. ***BMC Genomics*** 18 (Suppl 2):105
3. Chen, L and **Lu, X**. (2018) Discover functional impact of microRNAs in cancer using a causal deep learning model. ***BMC Medical Genomics,*** 11(Suppl 6):116.
4. Cai C, Cooper GF, Lu KN, Ma X, Xu S, Zhao Z, Chen X, Xue Y, Lee AV, Clark N, Chen V, Lu S, Chen L, Yu L, Hochheiser HS, Jiang X, Wang QJ, **Lu X**. (2019) Systematic discovery of the functional Impact of somatic genome alterations in individual tumors through tumor-specific causal inference. ***PLoS Computational Biology***.

**B. Positions and Honors**

1984-1985 Residency, Internal Medicine, Shengli Central Hospital, Dongying, China

1988-1991 Chief Resident, Dept. of Emergency Medicine, Shandong Provincial Hospital, Jinan, China

1991-1993 Attending cardiologist, Dept of Emerg. Med., Shandong Provincial Hospital, Jinan, China

1998-2001 Research Associate, Dept. of Pharmacology, University of Pittsburgh

2001-2003 National Library of Medicine training fellow, Center for Biomedical Informatics, University of Pittsburgh

2003-2008 Assistant Professor, Dept of Biostatistics, Bioinformatics and Epidemiology, Medical Univ. of South Carolina

2006-2010 Director, NLM training program “Training of Toolmakers for Bio-Medical Informatics”, Dept of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina

2007-2010 Director, Bioinformatics Division, Dept of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina

2008-2010 Associate Professor, Bioinformatics, Dept Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina

2008-2010 Co-Directors, DOE GAANN training grant and NIGMS T32 training grant

2009-2010 Associate Professor, Bioinformatics, Dept Biochemistry and Molecular Biology, Medical University of South Carolina

2010- 2016 Associate Professor, Biomedical Informatics, Department of Biomedical Informatics, University of Pittsburgh

2016 - Professor, Biomedical Informatics, Department of Biomedical Informatics, the University of Pittsburgh.

# Honors and Awards

# 1997 SmithKline Beecham Award for outstanding graduate student research at New England Pharmacologists’ Meeting. Boston, MA

# 1998 Fogarty Postdoctoral Fellowship Award, National Institutes of Health

# 2001-2003 National Library of Medicine Training Fellowship award

# 2003 Lister Hill National Center for Biomedical Communication Summer Research Participation Program Fellowship

# 2004 International Society for Computation Biology travel award for PSB 2004

# 2005 The Third International Charleston Ceramide Conference travel award

# 2009 Outstanding Paper Award, AMIA Summit on Translational Bioinformatics, San Francisco, CA

# Professional Services

# 2007-2016 Ad hoc member, multiple NLM Special Panel Study Section, BDMA,

# 2007-2014 Program Committee service: IEEE 7th International Symposium on Bioinformatics & Bioengineering (BIBE 2007); (BMEI 2008); ICMS(2011); AMIA (2011); WABI (2012); ICBC (2012); (ACBIT’2013); BioVis (2014), APBC 2016.

# 2008-2010 Member, NLM Study Section, Biomedical Library and Informatics Research Committee (BLIRC), CSR/NIH

# 2008 - Associate Editor, BMC Research Notes; Editorial Board, DNA Repair

2020 - Chair (2020 – 2021) and member (2020 – 2024) of CSR Study Section: Biomedical Informatics, Library and Data Sciences

# C. Contribution to Science

**C.1. Translational bioinformatics and computational cancer biology.** In recent years, my research concentrates on developing computational biology methodologies to study cancer signaling and disease mechanisms. We have designed gene expression module based models for predicting breast cancer patient survival (the [DREAM 7 Challenge](https://sagesynapse.wordpress.com/2012/11/01/breast-cancer-challenge-team-pitttransmed-places-second-for-metabric-phase-of-the-challenge/)) and our team won the 2nd best performance. Our algorithms for predicting signaling network has won the first place in the [SBV IMPROVER Trans-species Network Inference Challenge](https://sbvimprover.com/challenge-2/overview), and recently published in Cancer Informatics [19].

1. Cai, C., Chen, L., Jiang, X., and **Lu, X**., (2014) Modeling signal transduction from protein phosphorylation to gene expression. ***Cancer Informatics***, 13(S1):59-67
2. Huang, T., Alvarez, AA, Pangeni, RP., Horbinski, C., Lu, S., Kim, SK., James, CD., Raizer, J., Kessler, J., Brenann, CW., Sulman, EP., Finocchiaro, G., Tan, M., Nishikawa, R., **Lu, X**., Nakano, I., Hu1, B., and Cheng, SY.. (2016) A Regulatory circuit of miR-125b/miR-20b and Wnt signaling controls GBM phenotypes through FZD6-mediated pathways. ***Nature Communication.* 7**:12885
3. Chen, V., Paisley, J., and **Lu, X** (2017) Revealing common disease mechanisms shared by tumors of different tissues of origin through semantic representation of genomic alterations and topic modeling. ***BMC Genomics*** 18 (Suppl 2):105
4. Ding, MQ., Chen, L., Cooper, GF., Young, JD., and **Lu, X**. (2017) Precision oncology beyond targeted therapy: Combining omics data with machine learning matches the majority of cancer cells to effective therapeutics. ***Molecular Cancer Research*** 16(2):269-278
5. Huang, T, Kim, CK., Alvarez, AA., Pangeni, RP., Shi, T, Sastry, N., Lu, S., Horbinski, C., Kessler, J., Nishikawa, R., Nakano, I., **Lu, X**, James, CD., Ying, XM., Hu, B., and Cheng, SY. (2017) MST4 phosphorylation of ATG4B regulates autophagic activity, tumorigenicity and radioresistance in cancer. ***Cancer Cell*** 32:840-855

**C.2. Pharmacology and systems biology.** I have a long-standing interest in study cellular signaling transduction using experimental and computational approaches. In particular, working with biologist collaborators, my group has developed an integromic approach to study the signaling roles of a family of bioactive lipids known as sphingolipids. Due to highly interconnected metabolic network, studying signaling roles of individual sphingolipids has eluded conventional experimental studies for decades. Our work led to breakthroughs in the sphingolipid signaling domain, and our papers were published in high-impact journals like Molecular Systems Biology and Science Signaling [10, 12]. Recently, we designed *deep learning* model to perform trans-species learning, i.e., to predict human cell responses to certain stimuli based on the responses by rat cells. To our knowledge, this is the first paper [16] applying deep learning techniques to model cellular signaling system, which potentially will open a new research direction in bioinformatics.

1. Cowart, LA., Shotwell, M., Worley, ML., Richards, AJ, Montefusco, DJ, Hannun YA, and **Lu, X**.. (2010) Revealing a signaling role of PHS1P in yeast using integrative systems approaches. ***Molecular Systems Biology*** **6**:349 (Highlighted at ISMB 2010) [PMCID: PMC2835565](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2835565/)
2. Montefusco, D., Chen, L, Matmati, N., Lu, S., Newcomb, B., Cooper, GF., Hannun, YA., **Lu, X**., (2013) Distinct signaling roles of ceramide species in yeast revealed through systematic perturbation and systems biology analyses. ***Science Signaling*** 6:rs14 [PMCID: PMC3974757](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3974757/)
3. Osmanbeyoglu, H, Lu, K., Oesterreich, S, Day, RS, Benos, PV, Coronnello, C., and, **Lu, X** (2013) Estrogen represses gene expression through chromatin reconfiguration. ***Nucleic Acid Research*** 41(17): 8061-8071
4. Chen, L., Cai, C., Chen, V., and **Lu, X** (2015) Trans-species learning of cellular signaling systems with bimodal deep belief networks. ***Bioinformatics*** 31 (18): 3008-3015 (doi: 10.1093/bioinformatics/btv315)
5. Tao, Y., Cai, C., Cohen, W., and Lu, X (2020) From genome to phenome: Predicting multiple cancer phenotypes based on somatic genomic alterations via the genomic impact transformer. Proceedings of ***Pacific Symposium on Biocomputing.***

# C.3. Statistical text mining of biomedical literatures and automatic annotations. I have made significant contributions to text-mining in the bioinformatics domain. Our 2006 paper [1] is one of the first papers introducing PTMs to the bioinformatics field, which motivated many follow up studies of other groups, reflected by over 40 citations. Our efforts in modeling the semantic topics associated with genes/proteins further enabled us to assess if the functions of a set of genes are coherently related (functional coherence), automatic textual evidence identification and automatic function annotation of proteins, with numerous publications in Genome Biology, Bioinformatics and BMC Bioinformatics.

1. Zheng, B. and **Lu, X.** (2007) Novel metrics for evaluating the functional coherence of protein groups via protein-semantic-network. ***Genome Biology***, 8:R153 [PMCID:PMC2323239](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2323239/)
2. Jin, B., Muller, B., Zhai, CX, and **Lu, X** (2008) Multi-label literature classification based on the Gene Ontology graph.  ***BMC Bioinformatics*** 9:525 [PMCID: PMC2644325](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2644325/)
3. Jin, B. and **Lu, X** (2010). Identifying informative subsets of the Gene Ontology with information bottleneck methods. ***Bioinformatics*** **26** : 2445-2451 [PMCID: PMC2944202](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2944202/)
4. Chen, V., Paisley, J., and **Lu, X** (2017) Revealing common disease mechanisms shared by tumors of different tissues of origin through semantic representation of genomic alterations and topic modeling. ***BMC Genomics*** 18 (Suppl 2):105
5. Jin, Q., Dhingra, B., Cohen, W., and Lu, X. (2018) AttentionMeSH: Simple, Effective and Interpretable Automatic MeSH Indexer. ***Proc of Empirical Methods in NLP.*** Brussel, Belgium.

# D. Research Support

**Active**

**R01 LM012011 (Lu)**  05/2020 – 04/2024

Sponsor: NIH/NLM $1,238,488

**Title of Project: Interpretable deep learning models for translational medicine**

This project aims to develop explainable AI models for discovering signaling pathways underlying physiological and pathological processes of cancer cells. The inferred pathway-activation information will be used to predict cancer cell responses to anti-cancer drugs in precision oncology setting.

**R01 CA215481-01A1 (Yu)** 01/01/2018 – 12/31/2022

Sponsor: NIH $2,961,215

**Title of Project: Translation addiction and targeting in colon cancer**

The proposed work will develop and employ innovative tools, including phosphorylation defective eiF4E cancer cells and mice, clinical samples, and a highly mechanistic approach to define the role of p-eIF4E in colon cancer biology and therapy. If successful, our work will provide a better understanding on driver-induced prosurvival ER stress and metabolic adaptation, as well as strategies to target this addiction by promoting catastrophic ER stress and cell death. These findings are likely to have important implications in treating other cancers with hyperactive Wnt, for which no effective therapy currently exists.

**R01 CA229431-01A1 (Wang)** 07/02/2019 – 06/30/2024

Sponsor: NIH $1,552,056

**Title of Project**: **A novel mitotic regulatory axis in neuroendocrine prostate cancer**

Our study will define the role of PKD in mitotic regulation and contribute to the understanding of molecular mechanisms underlying TR-NEPC progression in late-stage mCRPC and uncover novel therapeutic strategies to treat this aggressive cancer.

**Completed (in last 4 years)**

R01LM011155 (PI: **Lu, X**) 09/2011—08/2015

NIH/NLM

**Ontology-driven methods for knowledge acquisition and knowledge discoveries**

This project develops probabilistic language models for mining biomedical literature, automatic annotating biomedical texts, and knowledge representation.

U54HG008540 (PI: Cooper, Bahar) 10/01/14-06/30/18

NIH/NHGRI

**Center for Causal Modeling and Discovery of Biomedical Knowledge from Big Data**

This National Center of Excellence in BD2K concentrates on developing, implementing, and evaluating an integrated set of tools that support causal modeling and discovery of biomedical knowledge from very large and complex biomedical data.

R01LM012011 (PI: **Lu**) 04/15 – 03/19

NIH/NLM

**Deciphering cellular signaling system by deep mining a comprehensive genomic compendium.**

This project develops deep learning models to reveal major cellular signals that regulate gene expression under physiological and pathological conditions and to infer the organization of signals in human cellular signal transduction systems.

P30CA047904 Pilot Project (PI: **Lu**) 05/18 – 04/19

**Developing an AI-based Clinical Decision Support System for Precision Oncology**

This project develops novel deep learning models to learn the state of cancer cell signaling systems and use such information to predict cancer cell sensitivity to existing anti-cancer drugs.

UPMC Immune Transplant and Therapy Center. IPA 2018 NO. 6 (PIs: **Lu**, Zarour). 1/18 – 12/19

**Tumor-specific causal inference for guiding immune therapy ($1.7M)**

This project was supported by the Immune Transplant and Therapy Center (a $200M initiative by the UPMC), The project retrospectively collected genomic data (whole exome sequencing and bulk RNA sequencing) and clinical data of 260 melanoma patients treated with PD-1 checkpoint inhibitor. The goal was to create commercial product using causal discovery methods to reveal distinct immune evasion mechanisms exploited tumors and use such information to predict sensitivity to current immune therapy.