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, Co-Director, Center for Translational Bioinformatics

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 the years of "wet-lab" research as a pharmacologist and the years in "dry-lab" using computational approaches. My current research concentrates on developing computational methods for studying cancer pathways and cancer pharmacogenomics. My group develops probabilistic models to study cellular signaling through mining genome-scale data and biomedical literatures, including developing latent variable models to represent the state of cellular signaling system, identifying transcription factor modules, modeling cellular signaling systems using **deep learning** algorithms. With a unique combination of expertise in molecular biology, cell biology, pharmacology, signal transduction pathway modeling, and machine learning, my lab is particularly suitable 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; our Bayesian network-based model for studying cellular signaling pathway won the first place in the SBV IMPROVER Trans-species Network Inference Challenge (https://sbvimprover.com/challenge-2/overview). As the contact PIs of this project, I will also work closely with Co-PI, Dr. Gregory Cooper, a world class researcher in causal discovery, to translate biological problems into computational ones and to design novel cancer driver and pathway discovery algorithms as we did past. In addition, I will be responsible for overall project, including participating design cell biology and pharmacological experiments, data analysis, knowledgebase design, writing manuscripts and reports to NIH.

Relevant publications on modeling signaling pathways:

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
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: PMC3974757Chen, V., Paisley, J., and Lu, X (2016) 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 PMC5374647
3. Lu, S., and **Lu, X** (2013) Using graph model to find transcription factor modules: the hitting set problem and an exact algorithm. ***Algorithms for Molecular Biology* 8:**2 PMCID: PMC3622577
4. 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
5. 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

6. Lu, S., Mandava, G., Yan, G., and **Lu, X** (2016) An exact algorithm for finding cancer driver somatic genome alterations: the weighted mutually exclusive maximum set cover problem. ***Algorithm for Molecular Biology*** **11**:1, DOI: 10.1186/s13015-016-0073-9

B. Positions

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
1998-1998	Postdoctoral Fellow, Signal Transduction Laboratory, National Institute of Environmental Health Sciences
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
2008-2009	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-	Associate Professor, Biomedical Informatics, Department of Biomedical Informatics, University of Pittsburgh
2010-	Co-Director, the Center for Translational Bioinformatics, School of Medicine, the University of Pittsburgh
2016 -	Professor, Biomedical Informatics, Department of Biomedical Informatics, the University of Pittsburgh.

Honors and Awards

1992	Best Teaching Award, Shandong Provincial Hospital Nursing School
1993-1998	Basic Medical Science Fellowship, University of Connecticut Health Center
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); Chair, Special Session on Statistical Methods in Biomarker Identification (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;
2012-	Editorial Board, DNA Repair
2016 -	Editorial Board, Scientific Reports

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) 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, and recently published in Cancer Informatics [19].

1. Feng, H, Hu, B, Liu, KW, **Lu, X**, Yiin, JJ, Lu, S, Keezer, S, Fenton, T, Furnari, FB, Hamilton, RL, Vuori, K, Nagane, M, Nishikawa, R, Cavenee, WK and Cheng, SY (2012) Activation of Rac1 by Src-dependent phosphorylation of Dock180Y1811 mediates PDGFRα-stimulated glioma tumorigenesis in mice and humans. ***Journal of Clinical Investigation*** 121(12):4670–4684
2. 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
3. Jiang X, Cai B, Xue D, **Lu X**, Neapolitan RE, Cooper GF (2014). A Comparative analysis of methods for predicting clinical outcomes using high-dimensional genomic datasets. ***J Amer. Med Info Assoc.*** **21**:e312-e319
4. 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
5. 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
6. Jonathan, YD, Cai, C., and Lu, X (2017) Unsupervised Deep Learning Reveals Prognostically Relevant Subtypes of Glioblastoma. ***BMC Bioinformatics*** (to appear)

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.

7. Lu, S., and **Lu, X** (2013) Using graph model to find transcription factor modules: the hitting set problem and an exact algorithm. ***Algorithms for Molecular Biology*** **8:**2 PMCID: PMC3622577

8. Qin, T., Tsoi, LC., Sims KJ, **Lu, X** and Zheng, WJ (2012) Signaling network prediction using the ontology fingerprint enhanced Bayesian networks. ***BMC Systems Biology* 6** (Suppl 3) : S3 (co-corresponding author) PMCID:PMC3524013
9. Richards, A., Schwacke, J., Rohrer, B., Cowart, LA. and **Lu, X** (2012) Revealing functionally coherent gene subset using spectral clustering and information integration approaches. ***BMC Systems Biology* 6** (Suppl 3) : S7
10. 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 PMC4668779
11. Chen, L., Cai, C., Chen, V., and **Lu, X** (2015) Learning a hierarchical representation of the yeast transcriptomic machinery using an autoencoder model. ***BMC Bioinformatics*** 17(Suppl 1):S9 PMC4895523.

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.

12. Zheng, B., McLean, DC., and **Lu, X**. (2006) Identifying biological concepts from a protein-related corpus with a probabilistic topic model. ***BMC Bioinformatics*** 7:58 PMID: 16466569
13. 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
14. 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
15. Jin, B. and **Lu, X** (2010). Identifying informative subsets of the Gene Ontology with information bottleneck methods. ***Bioinformatics* 26** : 2445-2451 PMCID: PMC2944202
16. Chen, V and **Lu, X** (2013) Conceptualization of molecular findings by mining gene annotations. ***BMC Proceedings* 7**(Suppl 7):S2
17. Ogoe, HA, Visweswaran, S, **Lu, X**, Gopalakrishnan, V. (2015) Knowledge transfer via classification rules using functional mapping for integrative modeling of gene expression data. ***BMC Bioinformatics*** 16:226 (designated as a Highly Accessed paper)

D. Research Support

Active

R01LM012011 (PI: **Lu**) 04/15 – 03/19
NIH/NLM
Deciphering cellular signaling system by deep mining a comprehensive genomic compendium.
In this project, we will compile a comprehensive compendium of human gene expression data from GEO and then employ modern deep-learning algorithms and supercomputers to mine the data. We aim 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.

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 Center of Excellence is 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.

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 computational tools to extract biological concepts from free text, to represent the knowledge for reasoning and discovery

1R01 LM 010144 (PI: **Lu, X,**) 07/2009 – 06/2014
NIH/NLM
Statistical methods for integromics discovery
This project aims to develop statistical methods for integrating information from multiple types of high-throughput biotechnologies for inferring bioactive-lipid-mediated signal transduction systems

2 R01 GM063265-09 (PI: Hannun, 0.6 months) 02/2000 – 02/2015
Modeling and Analysis of Roles of Yeast Sphingolipids
This project develops computational models to study yeast sphingolipid metabolism and signaling.

1R01GM100387 (PI: Gopalakrishnan, 0.6 months) 7/2013 – 6/2015
Transfer Rule Learning for Knowledge Based Biomarker Discovery and Predictive Biology.
10/2012 – 9/2015
This project develops rule-based approach to identify biomarkers.