OMB No. 0925-0001 and 0925-0002 (Rev. 10/2021 Approved Through 09/30/2024)

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: LUJIA CHEN

eRA COMMONS USER NAME: LUGA1012

POSITION TITLE: ASSISTANT PROFESSOR, BIOMEDICAL INFORMATICS

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE | Start Date | Completion Date | FIELD OF STUDY |
| --- | --- | --- | --- | --- |
| University of Science and Technology Beijing, Beijing, China  Medical University of South Carolina  South Carolina, USA | B.S.  M.S. | 08/2005  08/2009 | 07/2009  07/2010 | Biotechnology  Bioinformatics |
| University of Pittsburgh  PA, USA  University of Pittsburgh  PA, USA  University of Pittsburgh  PA, USA  University of Pittsburgh  PA, USA | M.S.  Ph.D.  Postdoc  Assistant Professor | 08/2010  08/2012  05/2016  02/2022 | 07/2012  04/2016  01/2022  Now | Biomedical Informatics  Biomedical Informatics  Biomedical Informatics  Biomedical Informatics |

|  |  |  |
| --- | --- | --- |
| **A. Personal Statement**  Dr. Chen is a tenure track assistant professor at the Department of Biomedical Informatics at the University of Pittsburgh. She has established a strong research background in bioinformatics, biomedical informatics, biology, and machine learning during her Ph.D. studies. She has been developing computational methods for studying the cellular signaling transduction system since 2010. She is especially proficient in bioinformatics and machine learning techniques. Her recent research concentrates on developing machine learning methods, especially deep learning models, to study cancer cell signaling systems, disease mechanisms, and pharmacogenomics. She is one of the first researchers who applied the deep learning model (DLM) to the biomedical field and published several papers showing the feasibility of DLM in analyzing biomedical data.  Dr. Chen received her K99/R00 award in 2019. The mentored phase (two-year K99 phase) of formal training has provided her with considerably more knowledge in cancer biology, pharmacology, Bayesian deep learning, and machine learning algorithms, which are major parts of her proposed training. These skills lay a foundation for her independent research phase (three-year R00 phase). The environment at University of Pittsburgh and UPMC provides her with enriched collaboration opportunities with experts from both cancer biology & immunology, and clinical experts from various diseases such as lung inflammation.  Relevant to this project, Dr. Chen has been concentrating on developing novel machine-learning models for modeling single-cell signaling and cell-cell communication for the last few years, and her work led to novel methodologies for modeling cell-cell communication at the individual tumor level, which lays a solid foundation for this study. She is applying the information derived from modeling cell signaling systems in the precision medicine domain.  Publications related to the proposed study.   1. *Chen, L.*, et al (2023) Machine learning predicts oxaliplatin benefit in early stage colon cancers. Journal of Clinical Oncology, in production. 2. Zhang, H., Lu, X, Lu, B, and *Chen, L.* (2023) scGEM: Unveiling the nested tree-structured gene co-expressing modules in single-cell transcriptome data. Cancers, 15(17), 4277. PMC9822106 3. Chen, X.\*, *Chen, L.\**, et al (2022) An individualized causal framework for learning intercellular communication networks that define microenvironments of individual tumors. PLOS Computational Biology, 18(12), p.e1010761. PMC9822106 4. *Chen, L.*, Cai C., Chen V., Lu X. (2016) Learning a hierarchical representation of the yeast transcriptomic machinery using an autoencoder model, BMC bioinformatics, 17 Suppl 1, 9. PMC4895523   **B. Positions and Honors**  **Positions**  02/2022-Now, Assistant Professor, Biomedical Informatics, Department of Biomedical Informatics, the University of Pittsburgh  11/2016-01/2022, Postdoc Fellow, Biomedical Informatics, Department of Biomedical Informatics, the University of Pittsburgh  01/2022 – present, tenure-track assistant professor, Department of Biomedical Inforamtics, University of Pittsburgh  **Honors and Awards** DREAM 8 CHALLENGES powered by Sage Bionetworks, 2nd place award, 2013  SBV IMPROVER 4, Trans-species Network Inference Challenge, 1st place award, 2014  **C. Contributions to Science**  **C.1 Analyzing single-cell RNA-sequencing data to study tumor microenvironment.** Cancer cells and cells within the tumor microenvironment together determine disease progression, tumor metastasis and response to/escape from immunotherapy. Dr. Chen developed an instance-specific causal analysis framework for discovering tumor-specific intercellular communication network (ICN) in the tumor microenvironment (TME) of HNSCC and CRC. The study shows that cellular states of cells in TMEs are coordinated through INCs that enable multi-way communications among epithelial, fibroblast, endothelial, and immune cells. The tumor microenvironment in tumor cells and normal cells are significantly different from each other to lead to tumor progression. The individual ICNs could lead to the discovery of novel different subtypes of networks that underlie disparate TMEs of tumors. Personalized treatment could be guided based on these novel subtypes. The built network from SGA to immune cells indicates the possible influence of somatic mutation/alteration on epithelial cells and more complex communication among immune cells. Besides, she used causal inference to study the functional impact of somatic genome alterations in individual tumors to improve the tumor-specific precision medicine.   1. Cai, C., Cooper, G.F., Lu, K.N., Ma, X., Xu, S., Zhao, Z., Chen, X., Xue, Y., Lee, A.V., Clark, N., Chen, V., Lu, S., *Chen, L.*, Lu X. (2019) Systematic Discovery of the Functional Impact of Somatic Genome Alterations in Individual Tumors through Tumor-specific Causal Inference, PLoS Comput Biol. PMC6650088 2. Chen, X., *Chen, L.*, et al (2022) An individualized causal framework for learning intercellular communication networks that define microenvironments of individual tumors. PLOS Computational Biology, 18(12), p.e1010761. PMC9822106 3. Zhang, H., Lu, X, Lu, B, and *Chen, L.(2023)* scGEM: Unveiling the nested tree-structured gene co-expressing modules in single-cell transcriptome data. Cancers, 15(17), 4277.   **C.2 Innovative Deep Learning Models in Biomedicine.**Developed from [artificial neural networks](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/artificial-neural-networks), **deep learning**-based algorithms show great promise in extracting features and learning patterns from complex data. Dr. Chen’s research concentrates on developing machine learning methods, especially deep learning models, to study cancer cell signaling systems and disease mechanisms. The key advantage of DLMs is their capability of using hierarchically organized latent variables to capture the context-specific statistical structure underlying the observed data. Dr. Chen is one of the first researchers who applied the deep learning in biomedical field. During the last few years, Dr. Chen has contributed to several innovative deep learning models specifically designed to solve biomedical problems. She designed a Bimodal DLM to transfer the knowledge learned from rat protein phosphorylation data to predict human protein phosphorylation. She used deep neural network with sparse regularization to model and investigate the hierarchical structure of the yeast signaling pathway and map hidden units in the model to real biological entities. She designed DLMs to learn hierarchical representations of cellular signaling pathways in cancer cell lines and use the concise representation learned from the DLMs to train machine learning classifiers to predict drug sensitivity. She designed DLMs allowing the regulatory edge between miRNAs and mRNAs to learn the regulatory mechanism between miRNA and mRNA, which could guide the identification of miRNA targets/biomarkers to predict the clinical outcomes. All the work listed above contributes to the cancer ontology and promotes the development of precision medicine.   1. *Chen, L.*, Cai C., Chen V., Lu X.(2015) Trans-species learning of cellular signaling systems with bimodal deep belief networks, *Bioinformatics*, **31**, 3008-3015. PMC4668779 2. *Chen, L.*, Cai C., Chen V., Lu X. (2016) Learning a hierarchical representation of the yeast transcriptomic machinery using an autoencoder model, BMC bioinformatics, 17 Suppl 1, 9. PMC4895523 3. Ding, M.Q., *Chen, L.*, Cooper G.F., Young, J.D., Lu X. (2018) Precision Oncology beyond Targeted Therapy: Combining Omics Data with Machine Learning Matches the Majority of Cancer Cells to Effective Therapeutics, Molecular cancer research : MCR, 16, 269-278. PMC5821274 4. Chen, L., et al (2023) Machine learning predicts oxaliplatin benefit in early stage colon cancers. Journal of Clinical Oncology, to appear.   **C.3 Studying signaling transduction system.** Cancer is among the leading causes of death that is associated with alteration of multiple cellular signaling pathways and biological processes in the cell. Therefore, signaling networks are of great importance for us to understand the cell's regulatory mechanism. Dr. Chen mainly used integrative analysis of multi-omics data to infer the signaling pathway. She used the manipulated perturbation data and transcriptomic data to infer the causal relationships between ceramide species and their potential targets by combining lipidomic, genomic, and transcriptomic analyses. She investigated computational methods that integrated proteomics and transcriptomic data to identify signaling pathways transmitting signals in response to specific stimuli. She modeled and inferred the perturbed signal transduction system from protein phosphorylation to gene expression. She leaned the transcriptomic signatures of p53 mutations by performing signal-oriented pathway analysis. She used a de novo deep learning method to search for yeast signaling pathways using genetic perturbation data.   1. Montefusco, D.J.\*, *Chen, L*.\*, Matmati N, Lu S, Newcomb B, Cooper G.F., Hannun Y.A., Lu X. (2013) Distinct signaling roles of ceramide species in yeast revealed through systematic perturbation and systems biology analyses, Science signaling, 6, rs14. (\* indicates co-first author) PMC3974757 2. Cai, C.*, Chen, L.,* Jiang, X., Lu X. (2014) Modeling signal transduction from protein phosphorylation to gene expression, *Cancer informatics*, **13**, 59-67. PMC4216050 3. Hill, S.M.*, et al.* (2016) Inferring causal molecular networks: empirical assessment through a community-based effort, *Nat Methods*, **13**, 310-318. PMC4854847 4. Lu, S., Cai C., Yan, G., Zhou, Z., Wan, Y., Chen, V., *Chen, L.*, Cooper GF., Oveid LM., Hannun Y.A., Lee A.V., Lu, X.(2016) Signal-Oriented Pathway Analyses Reveal a Signaling Complex as a Synthetic Lethal Target for p53 Mutations, *Cancer research*, **76**, 6785-6794. PMC5165695 5. Lu, S., Fan, X., *Chen, L.*, Lu, X. (2018) A novel method of using Deep Belief Networks and genetic perturbation data to search for yeast signaling pathways, *Plos One*, **13**, 9.   **C.4 Text Annotation.** Most of the knowledge regarding genes and proteins is stored in biomedical literature as free text. Extracting information from complex biomedical texts demands techniques capable of inferring biological concepts from local text regions and mapping them to controlled vocabularies. In this study, scLDA model is built to predict GO annotations by identifying text regions. Dr. Chen helped conduct the evaluation experiment.   1. Jin, B., Chen, V., *Chen, L.*, Lu, X. (2011) Mapping annotations with textual evidence using an scLDA model, AMIA ... Annual Symposium proceedings. AMIA Symposium, 2011, 834-842. PMC3243146   **Complete List of Published Work:**  <https://www.ncbi.nlm.nih.gov/myncbi/lujia.chen.1/bibliography/public/> |  |  |