, *ORCID*

xl7 at psc.edu & (202) 567-8999

Xiang(Shawn) Li, Ph.D.

INTERESTS

AI-Agentic System for Precision Medicine

EDUCATION

Ph.D., George Washington University, Physics (**Honor: Parke Prize**), 2016 ~ 2021
Thesis Summary Study of Transcriptional and Post-transcriptional Regulation in Immune Systems
B.S., Huazhong University of Science & Technology, Applied Physics, (**Honor**) 2011 ~ 2015

Work Experiences

PSC, Carnegie Mellon University Senior Scientist-AI-Agent Lead Pittsburgh PA, 2024 ~ Present

- **Partnership** Human BioMolecular Atlas Program: Lead the collaboration between TACC & PSU Galaxy, By building AI-enabled research infrastructure and workflows(HuBMAP) under Common Fund Data Ecosystem (CFDE)
- **Platform Built**: AI-assisted data research + multi-agent analytics infrastructure.
- **What Enabled**: Faster internal consulting response cycles; scalable scientific synthesis and analysis execution for multi-omics and biomedical tasks.
- **Scale Supported**: Multi-user environment; reusable agent workflows; integrated services across research users and consortium use-cases.
- **Lead** a multi-agent platform(MCP server + LangGraph/deepagents architecture), enabling enabling distributed scientific querying, evidence retrieval, and bioinformatic analysis execution.

Johnson & Johnson Scientist III, RWE, End due to working VISA(H1B). Titusville NJ, 2024

- R&D in Real World Evidence, Patient profiling from medical history (EHR, EMR)
- Dev a toolset from scratch for patient profile, alignment, visualization and analysis(built a fully connected network for 52K patients).

GentiBio, Senior Scientist, Bioinformatics Lead, End due to Reorg Boston MA, 2022 ~ 2024

- **Translational Research**: GNTI-122 IND package (Safety assessment for Gene Editing).
  Build SNIP-aware (gnomAD v4.0) insilico Off-target analysis for Cas9 editing
  On/Off-Target analysis protocol: in-Silico ⇒GUIDESeq ⇒rhAMPSeq.
  Build a full-stack web app (aws) for automation. Similar to CRISPResso2 Web
- **Early Discovery**, Scientist II: Cellular Profiling Platform
  Research: single cell multi-omics integration analysis
  Dev: AWS VC-backed start-up package and build bioinformatics infrastructure

Sana Biotechnology Scientist, Computational Biology, Seattle WA, 2021 ~ 2022

- **Corporate team**, research on target discovery & dev pipelines automation.
- **iPSC Master Cell Bank**, Cell Characterization & Safety Assessment

---

GWU, Department of Physics, Washington, DC, 2017 ~ 2021
Ph.D, (Superviser: *Prof. W. Peng* , Collaborator: *Prof. H. Xue*)

- *Nature Com:* Develop analysis framework (python-base for data representation, visualization and ML)
- Integration of Multi-omics: scRNAseq(10X), RNAseq and DNase_seq, ATAC_seq, ChIPseq, Cut&Run, HiC, MiCroC, HiChIP and etc.

- Chromatin structure and function: Develop a tool (HiCHub, which solved the problem on identifying a mixed range of altered genome structure(from looping to larger than TAD). This tool outperforms similar software in HiC data analysis with a significantly increase in the computational efficiency.
- Nucleic Acids Research: mRNA stability Mathematical Modeling and data analysis.
- Study regulatory roles of Tcf1/Lef1 in $CD8^+$ T cells in immunue response. (JEM, 2019, 2018)
- Impact upon deficiency of miRNA-34b/c during somatic reprogramming progress.
- Organize Journal Club and give presentation once per month.
- HPC management. (Linux Admin, data storage &sharing& backup, computation dependencies setup, jupyter notebook & Rstudio server and on-call for tech emergency.
- Build analysis pipeline, dependencies and visualization tools on a cloud platform. (AWS, Linux ssh).

Team Lead, NLP Course Project
George Washington University — Washington, DC, 2019

- Natural language processing, use different models for user profiling task in given posts from twitter.
- Write model to implement different language models, Hidden Markov model and bayesian classifier.

Joint Leader of Student Innovation Group (Adviser: B. Yu)
Qiming College of Huazhong University of Science & Technology — Boston, 2014

- Built a team for 2014 *BIOMOD, Harvard*, Cambridge, Massachusetts.
- Invested nano-scale transportation system based with magnetic control device.
- Conducted the final presentation on behalf of our team.
- For more detail, please visit our website: *BIOMOD 2014: HUST*

Teaching

- Graduate TA. General Physics (with: *X. Qiu* / *G. Younes* / *S. Guiriec*) — 2016 ~ 2018
- Instructor. Astrophysics (with: *Kalvir Dhuga*) — 2019

Membership & Reviewer

- Phenomics, Youth Editorial Board Members — 2024 ~ Present
- Society of Artificial Intelligence Research SAIR — 2023 ~ Present
- Machine Learning for Genomics MLGenX — 2024

Honors and Awards

- The Parke Prize, awarded for excellence in theoretical biophysics
- Scholarship in Department of Biostatistics, University of Washington, $1,950
- Second Prize of Poster in 3rd International Ocean Sciences Summer School. $800
- Project Awards: Silver; International Biomolecular Design Competition 2014, *BIOMOD, Harvard*
- Fellowship $12000 total, Undergraduate Fellowship from College of Life Science & Technology,

Publications & Patents & Presentations

- Inhibition of LARP4-mediated quiescence exit of naive CD4+ T cells ameliorates autoimmune and allergic diseases Jian Zhou, Di Yang, Chao Han, Hui Dong, Shufeng Wang, **Xiang Li**, ..., Yi Tian†., Nature Biomedical Engineering (2025)
- HuBMAP Data Portal: A Resource for Multi-Modal Spatial and Single-Cell Data of Healthy Human Tissues Jian Zhou, Di Yang, Chao Han, Hui Dong, Shufeng Wang, **Xiang Li**, ..., Yi Tian†., Nature Communications (NCOMMS-25-81343)
- HiCHub: A Network-Based Approach to Identify Domains of Differential Interactions from 3D Genome Data **Xiang Li**†, ..., Hai-Hui Xue, Weiqun Peng., Bioinformatics (BIOINF-2025-0207)
- Lysine methyltransferase Kmt2d regulates naive CD8+ T cell activation-induced survival Jaekwan Kim, T.N, ..., Xiang Li, ..., Nan-ping Weng., Frontiers in Immunology (2023)

- Tcf1 CTCF cooperativity shapes genomic architecture to promote CD8+ T cell homeostasis Q. S[†], ..., **Xiang Li**,..., Hai-Hui Xue., Nature Immunology (2022)
- Tcf1 and Lef1 maintain CD8+ T cell identity by organizing genomic architecture Q. S[†], **Xiang Li**[†],..., Hai-Hui Xue., Nature Communications (2021)
- **Selected Oral Presentation**, 2021 Network Biology, Cold Spring Harbour Laboratory Network analysis reveal the critical role of TCF1/LEF1, Speaker: **Xiang Li**
- Sepsis leads to lasting changes in phenotype and function of memory CD8 T cells Isaac J. Jensen[†], **Xiang Li**[†],..., Weiqun Peng and Vladimir P. Badovinac., 2021 eLife
- Arid1a-Plagl1-Hh signaling is indispensable for differentiation-associated cell cycle arrest of tooth root progenitors Yang Chai, ..., **Li Xiang**, Xinquan Jiang. Cell Reports (2021).
- Transcriptome-wide stability analysis uncovers LARP4-mediated NFB1 mRNA stabilization during T cell activation. Tian Y.[†], Zeng Z.[†], **Li Xiang**[†], ..., Peng W. and Zhu J. Nucleic Acids Research (2020).
- Integrated analysis of carbon dioxide and oxygen concentrations as a quality control of ocean float data Yingxu Wu, D. Bakker, E. Achterberg, ..., **Xiang Li**, ... Communications Earth Environment, Accepted. (Mar, 2022)
- Tcf1 and Lef1 transcription factors are required for the immunosuppressive function of regulatory T cells Xing S., Gai K., **Li Xiang**,..., Peng W. and Xue H., J Exp Med (2019).
- Tle corepressors are differentially partitioned to instruct CD8+ T cell lineage choice and identity Xing S., ..., **Xiang Li**,..., Peng W. and Xue H., J Exp Med (2018).

Work extensively used HiCHub

- Tcf1 CTCF cooperativity shapes genomic architecture to promote CD8+ T cell homeostasis Nature Immunology (2022)
- CTCF mediates CD8+ effector differentiation through dynamic redistribution and genomic reorganization JEM (2023)

To:
Search Committee & Dr. Jake Chen
Department of Biomedical Informatics and Data Science
School of Medicine at UAB

Cover Letter For Members of the Search Committee

I am writing to apply for a faculty position in the Department of Biomedical Informatics and Data Science at the School of Medicine at UAB. My research sits at the intersection of deep learning, agentic systems, and biomedical multi-omics modeling for gene regulation, with a central focus on developing agentic AI systems that transform biomedical data into actionable knowledge for precision medicine.

I am currently a Senior Scientist and AI-Agent Lead at the Pittsburgh Supercomputing Center and Carnegie Mellon University, where I lead the development of AI-enabled research infrastructure for large-scale biomedical data ecosystems, including the Human BioMolecular Atlas Program (HuBMAP). In this role, I designed and deployed a multi-agent analytics platform that integrates distributed querying, evidence retrieval, and bioinformatics execution. This system enables scalable, reproducible scientific discovery across multi-omics datasets and supports collaborative research in a multi-user environment. My work directly addresses core challenges in biomedical informatics: data integration, knowledge representation, and the operationalization of complex analytical workflows.

My prior experience spans translational informatics, clinical data science, and early discovery. At Johnson & Johnson, I worked in real-world evidence (RWE), developing patient profiling systems from electronic health records and constructing network-based representations of over 50,000 patients for clinical insights. At GentiBio and Sana Biotechnology, I led computational efforts in gene editing and cell therapy, including building end-to-end pipelines for CRISPR off-target analysis and multi-omics integration for cellular profiling. These experiences have shaped my ability to bridge computational methods with clinically relevant applications.

During my Ph.D. in Physics at George Washington University, I developed computational frameworks for integrating diverse genomic and epigenomic data modalities, including single-cell RNA-seq, ATAC-seq, and Hi-C. I built HiCHub, a network-based method for detecting differential chromatin interactions, which significantly improved computational efficiency and has been widely adopted in a series of immune cell studies. My publication record includes work in leading journals such as *Nature Communications* and *Nature Immunology*, reflecting contributions to both methodological innovation and biological discovery.

Looking forward, my research program aims to establish a new paradigm for biomedical discovery through agentic systems. Specifically, I plan to develop autonomous, multi-agent platforms that can
(1) continuously ingest and harmonize heterogeneous biomedical data,
(2) generate and test hypotheses through integrated modeling and simulation,
(3) provide interpretable, evidence-backed insights for precision medicine. These systems will function as collaborative scientific partners, accelerating early discovery pipelines and enabling more efficient translation from data to therapeutics.

I am particularly excited about the opportunity to contribute to UAB's interdisciplinary environment, where expertise across bioinformatics, clinical informatics, and translational science creates a strong foundation for impactful research. I am also committed to mentoring trainees and contributing to educational programs in biomedical data science and AI-driven discovery.

Thank you for your consideration. I would welcome the opportunity to discuss how my research and vision align with the goals of your department.

Sincerely,

Xiang Li, Ph.D.
03/25/2026