

## Xiang(Shawn) Li, Ph.D.

### INTERESTS

NGS Development, Algorithm & Machine Learning

### EDUCATION

Ph.D., George Washington University, Biophysics (**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

---

### SKILLS

Certificate: [Web App: WhenGreen](#), [AWS CCP](#), [DeepLearning.AI](#), [Nextflow Certificate](#),  
*Program, Software Skills and Database:*

- Advanced: Linux Admin, Shell Script, *Python*(data mining→visual→ML), NGS Tools,
- Intermediate: PyPi, Nextflow, S3 & EC2, flask, dash, boto3, Fortran, Gitlab, Github, SGE,  $\LaTeX$
- Basic: Snakemake, Docker, javascript, Web and Server Support(Nginx), MATLAB, C/C++, Java, R
- DB: genomAD, 4DN, NCBI, GEO, ENCODE, HapMap, 1000 Genomes, GIAB, dbSNP

---

### Work Experiences

[GentiBio](#), Senior Scientist, Bioinformatics Lead in a small team Boston MA, 2022 ~ Present

- Build computational components for GENTI122 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  $\Rightarrow$  *GUIDESeq*  $\Rightarrow$  *rhAMPSeq*.  
Build a full-stack web app (aws) for automation. Similar to [CRISPResso2 Web](#)
- Build Cellular Profiling Platform:  
scRNAseq & scATACseq integration analysis, Kidney biopsies from Mouse IgA Model  
Deploy GentiBio internal chatGPT & Automation pipeline through AWS  
Deploy AWS VC-backed startup package and build Bioinformatics Infrastructure.  
Write analysis tools (python & R) and pipeline (bash, Nextflow Docker)  
Adopt bulk/scRNAseq analysis pipeline and downstream analysis  
CAR design, Implement AlphaFold2 & AlphaMissense for IBD, FliC CAR.

[Sana Biotechnology](#) Scientist, Computational Biology Seattle WA, 2021 ~ 2022

- Work as IC in corporate team, focusing on computational pipelines (nextflow) development.
- iPSC Master Cell Bank, Cell Characterization & Safety Assessment
- GUI tools develop for analytical team

[Guardant Health](#) Intern, Scientist at Algorithm Team Redwood, CA, 2021

- 1. Study tumor fraction prediction by cell-free DNA fragment size distribution. (\*Built a Deep Learning module for production pipeline(BIP4.0).
  - 2.Study of a rare contamination found in sequence. \*Support variants calling by wrote a python based filtering module for BIP4.0.
-

Department of Physics, [GWU](#)

Washington, DC, 2017 ~ 2023

Ph.D, (Supervisor: [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<sup>+</sup> 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

## Consultant

- Computational Workflow Development for [Verily Life Sciences](#) 2023 ~ Present
- Computational Workflow Development for Insights Driven Research (IDR) one time

## Reviewer

- Machine Learning for Genomics [MLGenX](#) 2024
- [Phenomics](#) 2024 ~ Present
- Society of Artificial Intelligence Research [SAIR](#) 2023 ~ Present

## HONORS AND AWARDS

- The Parke Prize, awarded for excellence in theoretical biophysics
- [Visiting Scholar](#) (Host: Chairman Emeritus of CME Group [Leo Melamed](#) with [Dr. Z. Jin](#))
- 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

- [HiCHub: A Network-Based Approach to Identify Domains of Differential Interactions from 3D Genome Data](#) **Xiang Li**<sup>†</sup>, ..., Hai-Hui Xue, Weiqun Peng., bioRxiv (TBD)
- [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<sup>†</sup>, ..., **Xiang Li**,..., Hai-Hui Xue., *Nature Immunology* (2022)
- [Tcf1 and Lef1 maintain CD8+ T cell identity by organizing genomic architecture](#) Q. S<sup>†</sup>, **Xiang Li**<sup>†</sup>,..., 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<sup>†</sup>, **Xiang Li**<sup>†</sup>,..., 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.<sup>†</sup>, Zeng Z.<sup>†</sup>, **Li Xiang**<sup>†</sup>, ..., 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)