Xin Dong

Ph.D. in Bioinformatics
Department of Bioinformatics, School of Life Sciences and Technology, Tongji University
NO.1239 Siping Road, Shanghai, China
Website: https://xindong95.github.io/
Email: dongxin@tongji.edu.cn; xindong9511@gmail.com; **Phone**: +86 176 2166 5585

EDUCATION

Sep. 2018 – Jul. 2024	**Ph.D. in Bioinformatics** School of Life Sciences and Technology, Tongji University, China Advisors: Dr. Chenfei Wang & Dr. Xiaole Shirley Liu
Sep. 2014 – Jun. 2018	**B.S. in Biotechnology** College of Life Sciences, Shandong Normal University, China

RESEARCH EXPERIENCE

Graduate Student, Tongji University (2018-2024)

- **Developed and applied computational methods to analyze single-cell and spatial data**
  Advisors: Dr. Chenfei Wang & Dr. Xiaole Shirley Liu
    - To search and screen single-cell transcriptome and chromatin accessibility datasets from the GEO of NCBI, I developed a parser that can automatically obtain the metadata and download the raw data. This provides the foundation for our subsequent works.
    - We developed SCRIP (**S**ingle-**C**ell **R**egulatory Network **I**nference using ChI**P**-seq and motif) (https://github.com/wanglabtongji/SCRIP), a method that integrates scATAC-seq and a large collection of bulk TR (Transcription Regulator) ChIP-seq to infer single-cell TR activity and targets. Our method outperformed other motif-based approaches in assessing TR binding activity. It also enabled us to identify TR target genes and construct GRNs (Gene Regulation Networks) at the single-cell level using a regulatory potential model. We applied SCRIP to various biological systems and showed its utility in lineage tracing, cell-type clustering, and inferring cell-type-specific GRNs.
    - We developed TISCH (**T**umor **I**mmune **S**ingle **C**ell **H**ub; http://tisch.comp-genomics.org) and TISCH2, a large curated resource that integrates the single-cell transcriptome profiles of about 6 million cells from 190 high-quality tumor datasets across 50 cancer types. TISCH allows systematic comparison of gene expression across multiple datasets at the single-cell or cluster level for various cell types, patients, tissue origins, treatment and response groups, and even cancer types. TISCH aims to address the computational challenges of using the massive published datasets to inform immunotherapy and to characterize the immune system heterogeneity in cancers.
- **Large-scale genomic data mining and integration analysis**
  Advisors: Dr. Xiaole Shirley Liu & Dr. Clifford Meyer
    - To download the raw sequencing data of the chromatin occupancy experiments from the NCBI's GEO database and extract the metadata, we created a parser. The parser annotated and labeled the experiment type, tissue type, cell population, and other details of the experiments using the keywords-match strategy.
    - We developed CHIPS (**CH**romatin enr**I**chment **P**roce**S**sor) (https://github.com/liulab-dfci/CHIPS), a pipeline that processes the epigenome data to readable results with a single command line. It integrates states-of-the-art tools for cistrome NGS data and chooses the appropriate analysis parameters. It uses the Snakemake framework, which makes it easy to deploy on cloud or cluster servers.
    - We developed and maintained the Cistrome Data Browser (http://cistrome.org/db/#/), which hosts many public cistrome sequence datasets. It offers a user-friendly interface to access the data quality and the results

of downstream analysis.

- I investigated the reads-keeping strategy of ATAC-seq that captures the chromatin accessibility of the genome. I discovered that using only the fragments with less than 150 bp length can yield more peaks from ATAC-seq. I implemented this strategy in the CHIPS pipeline.
- We used unsupervised clustering to identify the low-quality dataset and examine the characteristics of high-quality and low-quality data for broad signal histone modifications, such as H3K27me3, which are challenging to assess the quality of the experiment.

Undergraduate Student, Shandong Normal University (2014-2018)

- **Rapid establishment of H7N9 dominant lineage by intragenic recombination in HA segment**
  Undergraduate Dissertation Research; Advisor: Dr. Chengqiang He
- **Screening of newcastle disease virus heat-resistant vaccine by homologous recombination**
  Undergraduate Innovation Program (Presided); Advisor: Dr. Chengqiang He
- **The origin, adaptive evolution and population dynamics of infectious Chicken Anemia Virus in China**
  Undergraduate Innovation Program (Involved); Advisor: Dr. Chengqiang He
- **Diversity of fish in Nansi Lake**
  Undergraduate Innovation Program (Involved); Advisors: Dr. Rongshu Fu & Dr. Mingsheng Miao

Skills

- **Bioinformatics**: Single-cell data analysis; High-throughput sequencing analysis; Epigenome data analysis; Machine learning; Deep learning
- **Programming**: Python; R; Shell; JavaScript; MySQL
- **High performance computing**: Cloud computing; Cluster computing; Snakemake pipelines; Python package development
- **User interface development**: Django; Vue.js; Bootstrap
- **Code version control system**: git
- **Server management**: Ubuntu; CentOS / Rocky Linux
- **Experiment skills**: Cell culture; PCR; Gel electrophoresis

PUBLICATIONS

Asterisk (*) denotes the Co-first Authors and Hashtag (#) denotes the Corresponding Authors.

- Chang, Z.*, Xu, Y.*, **Dong, X.***, Gao, Y.# & Wang, C.# Single-cell and spatial multiomic inference of gene regulatory networks using SCRIPro. *Bioinformatics* 40, btae466 (2024).
- Sun, F.*, Li, H.*, Sun, D.*, Fu, S.*, Gu, L.*, Shao, X.*, Wang, Q.*, **Dong, X.***, Duan, B.*, Xing, F.*, Wu, J.*, Xiao, M.#, Zhao, F.#, Han, J.-D. J.#, Liu, Q.#, Fan, X.#, Li, C.#, Wang, C.# & Shi, T.# Single-cell omics: experimental workflow, data analyses and applications. *SCIENCE CHINA Life Sciences* (2024).
- Ren, P.*, Shi, X.*, Yu, Z., **Dong, X.**, Ding, X., Wang, J., Sun, L., Yan, Y., Hu, J., Zhang, P., Chen, Q., Zhang, J., Li, T. & Wang, C.# Single-cell assignment using multiple-adversarial domain adaptation network with large-scale references. *Cell Reports Methods*, 3, 100577 (2023).
- Han, Y.*, Wang, Y.*, **Dong, X.***, Sun, D., Liu, Z., Yue, J., Wang, H., Li, T.# & Wang, C.# TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Research*, 51, D1029-D1037 (2022).
- Shi, X.*, Yu, Z.*, Ren, P., **Dong, X.**, Ding, X., Song, J., Zhang, J., Li, T.# & Wang, C.# HUSCH: an integrated single-cell transcriptome atlas for human tissue gene expression visualization and analyses. *Nucleic Acids Research*, 51, D1425-D1431 (2022).
- **Dong, X.***, Tang, K.*, Xu, Y., Wei, H., Han, T. & Wang, C. # Single-cell gene regulation network inference by large-scale data integration. *Nucleic Acids Research*, 50, e126 (2022).

- Xu, R.*, Li, S.*, Wu, Q.*, Li, C.*, Jiang, M.*, Guo, L., Chen, M., Yang, L., **Dong, X.**, Wang, H., Wang, C.#, Liu, X.#, Ou, X.# & Gao, S.# Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. *Cell Stem Cell*, 29, 1051-1066.e8 (2022).
- Sun, D.*, Wang, J.*, Han, Y.*, **Dong, X.**, Ge, J., Zheng, R., Shi, X., Wang, B., Li, Z., Ren, P., Sun, L., Yan, Y., Zhang, P., Zhang, F.#, Li, T.# & Wang, C.# TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Research*, 49, D1420-D1430 (2021).
- Zheng, R.*, **Dong, X.***, Wan, C., Shi, X., Zhang, X.# & Meyer, C.A.# Cistrome Data Browser and Toolkit: analyzing human and mouse genomic data using compendia of ChIP-seq and chromatin accessibility data. *Quantitative Biology*, 8, 267-276 (2020).
- Chen, C.-H., Zheng, R., Tokheim, C., **Dong, X.**, Fan, J., Wan, C., Tang, Q., Brown, M., Liu, J.S., Meyer, C.A. # & Liu, X.S.# Determinants of transcription factor regulatory range. *Nature Communications*, 11, 2472 (2020).
- Li, S.*, Wan, C.*, Zheng, R., Fan, J., **Dong, X.**, Meyer, C.A.# & Liu, X.S.# Cistrome-GO: a web server for functional enrichment analysis of transcription factor ChIP-seq peaks. *Nucleic Acids Research*, 47, W206–W211 (2019).

CONFERENCES

2024 Poster Speaking (Awarded), The 2nd Bioinformatics Conference of Yangtze River Delta, Shanghai, China
2023 Poster Speaking, Systems biology of gene regulation and genome editing, Cold Spring Harbor Conferences Asia, Suzhou, China
2023 Poster Speaking, The WLA Research Conference on Cells and Genes, Shanghai, China
2023 Poster Speaking, The 11st National Conference on Bioinformatics and Systems Biology, Guangzhou, China
2023 Poster Speaking, Human Cell Atlas Asia 2021 Meeting, Online
2021 Poster Speaking, The 10th National Conference on Bioinformatics and Systems Biology, Chengdu, China
2019 Attendance, 14th International Bioinformatics Workshop, Beijing, China

TEACHING

2023 Teaching assistant, Summer Deep Learning Course, School of Life Sciences and Technology, Tongji University, Shanghai, China
2022 Teaching assistant, Summer Bioinformatics Course, School of Life Sciences and Technology, Tongji University, Shanghai, China
2021 Teaching assistant, Summer Bioinformatics Course, School of Life Sciences and Technology, Tongji University, Shanghai, China
2020 Teaching assistant, Summer Bioinformatics Course, School of Life Sciences and Technology, Tongji University, Shanghai, China
2019 Teaching assistant, Dragon Star Course, The 14th International Bioinformatics Workshop, Beijing, China

AWARDS AND DISTINCTIONS

2024 Outstanding Graduates Awards of Shanghai, Shanghai Municipal Education Commission
2023 The 17th Academic Pioneer, Tongji University
2023 National Scholarship, Ministry of Education of the People’s Republic of China
2022 Outstanding Student Leader, Tongji University
2019 Outstanding Graduate Student Scholarship, Tongji University
2018 Outstanding Undergraduate Thesis for Bachelor’s Degree, Shandong Normal University
2017 Outstanding Graduates Awards of Shandong, Human Resources and Social Security Department of Shandong Province
2017 The First Prize of the 15th Challenge Cup Competition in Shandong, Shandong Association for Science and Technology
2017 Outstanding Student Leader of Shandong, Shandong Provincial Education Department

2016 The Grand Prize of Shandong College Students' Biochemistry Experiment Skills Competition, Shandong Association of Science & Technology