

Xin Dong

Ph.D. Candidate

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.2020 – Present

Ph.D. Candidate in Bioinformatics

School of Life Sciences and Technology, Tongji University, China

Advisors: Dr. Chenfei Wang & Dr. Xiaole Shirley Liu

Sep.2018 – Aug.2020

M.S. Candidate in Bioinformatics

School of Life Sciences and Technology, Tongji University, China

Advisors: Dr. Xiaole Shirley Liu & Dr. Xiaoyan Zhang & Dr. Zhiping Weng

Sep.2014 – Jun.2018

B.S. in Biotechnology

College of Life Sciences, Shandong Normal University, China

Advisor: Dr. Chengqiang He

RESEARCH EXPERIENCE

Graduate Student, Tongji University (2018-Present)

- **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 modified parser. The parser can automatically obtain, label, and download the data. This provides the foundation for our subsequent work.
 - We developed SCRIP (Single-Cell Gene Regulation Inference using ChIP-seq) (<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 (Tumor Immune Single Cell Hub; <http://tisch.comp-genomics.org>), a large curated resource that integrates the single-cell transcriptome profiles of about 2 million cells from 76 high-quality tumor datasets across 27 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 enrichment **ProceS**sor) (<https://github.com/liulab-dfci/CHIPS>), a pipeline that processes the epigenome data to readable results with a single command line. It integrates state-of-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:** High-throughput sequence data analysis; Single-cell data analysis; Epigenome data analysis; Machine Learning and statistic modeling
- **Programming:** Python; R; shell; JavaScript; MySQL

- **High performance computing:** Snakemake pipeline development; Python package development; Cloud computation; Cluster computation
- **User interface development:** Django; Vue.js; Bootstrap
- **Code version control system:** git
- **Server Management:** Ubuntu; CentOS
- **Experiment skills:** Cell culture; PCR; Gel electrophoresis

PUBLICATIONS

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

- Han,Y.* , Wang,Y.* , **Dong,X.***, Sun,D., Liu,Z., Yue,J., Wang,H., Li,T.# , Wang,C.# (2022) TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Research*, 10.1093/nar/gkac959.
- Shi, X.* , Yu,Z.* , Ren,P., **Dong,X.**, Ding,X., Song,J., Zhang,J., Li,T.# , Wang,C.# (2022) HUSCH: an integrated single-cell transcriptome atlas for human tissue gene expression visualization and analyses. *Nucleic Acids Research*, 10.1093/nar/gkac1001.
- **Dong,X.***, Tang,K.* , Xu,Y., Wei,H., Han,T., Wang,C.# (2022) Single-cell gene regulation network inference by large-scale data integration. *Nucleic Acids Research*, 10.1093/nar/gkac819.
- 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.# (2022) Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. *Cell Stem Cell*, 29, 1051-1066.e8.
- 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.# (2021) TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Research*, 10.1093/nar/gkaa1020.
- 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.# (2020) Determinants of transcription factor regulatory range. *Nature Communications*, 11, 2472.
- Zheng,R.* , **Dong,X.***, Wan,C., Shi,X., Zhang,X.# and Meyer,C.A.# (2020) Cistrome Data Browser and Toolkit: analyzing human and mouse genomic data using compendia of ChIP-seq and chromatin accessibility data. *Quantitative Biology*
- Li,S.* , Wan,C.* , Zheng,R., Fan,J., **Dong,X.**, Meyer,C.A.# and Liu,X.S.# (2019) Cistrome-GO: a web server for functional enrichment analysis of transcription factor ChIP-seq peaks. *Nucleic Acids Research*, 47, W206–W211.

CONFERENCES

- 2023 Poster Speaking, 11st National Conference on Bioinformatics and Systems Biology, 2023, Guangzhou, China
- 2021 Poster Speaking, Human Cell Atlas Asia 2021 Meeting, 2021, Online
- 2021 Poster Speaking, 10th National Conference on Bioinformatics and Systems Biology, 2021, Chengdu, China
- 2019 Attendance, 14th International Bioinformatics Workshop, 2019, Beijing, China

TEACHING

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

AWARDS AND DISTINCTIONS

- 2022 Outstanding Student Leader, Tongji University
- 2022 Social Activities Scholarship, Tongji University
- 2019 Outstanding Graduate Student Scholarship, Tongji University
- 2018 Outstanding Dissertation for Bachelor's Degree, Shandong Normal University
- 2017 Outstanding Graduates Awards of Shandong Province, Human Resources and Social Department of Shandong Province
- 2017 The First Prize of The Fifteenth Challenge Cup Competition in Shandong, Shandong Youth League
- 2017 Outstanding Student Leader in Shandong, Shandong Provincial Education Department
- 2016 The Grand Prize of Shandong College Students' Biochemistry Experiment Skills Competition, Shandong Association of Science & Technology

LEADERSHIP AND PUBLIC SERVICE

- 2023** **Director, Chairman of Organizing Committee**
Graduation party of School of Life Sciences and Technology, Tongji University
- 2022** **Volunteer of PCR Test for COVID-19**
Liferiver Bio-Tech Corp., Shanghai, China
- 2020** **Census Taker**
7th Census of China
- 2016 – 2017** **Vice-president**
Chairman of Executive Committee
Students' Union, Shandong Normal University
- 2015 – 2016** **Minister of Literatures and Arts Department**
Students' Union, Shandong Normal University
- 2016** **Chairman of Organizing Committee**
30th Top Ten Campus Singers Competition of Shandong Normal University
- 2015** **Chairman of Organizing Committee**
Golden Microphone Host Competition of Shandong Normal University
- 2015 – 2016** **Committee Member of Network Department**
Students' Union, School of Life Sciences, Shandong Normal University
- 2014 – 2015** **Member of Rights and Interests Department**
Students' Union, Shandong Normal University