

# Xin Dong

Ph.D. Candidate

Department of Bioinformatics, School of Life Sciences and Technology

Tongji University, 1239 Siping Road, Shanghai, China

Website: <https://xindong95.github.io/>

[dongxin@tongji.edu.cn](mailto:dongxin@tongji.edu.cn); [xindong9511@gmail.com](mailto:xindong9511@gmail.com)

+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)

- **Developing Computational Methods for Single-cell and Spatial Data**

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

- We developed a modified parser to screen single-cell transcriptome and chromatin accessibility datasets from the GEO of NCBI. The data can be automatically obtained, labeled, and downloaded by the parser. It offers the building blocks for the works that follow.
- Single-cell ATAC-seq is an approach to investigating gene regulation at the single-cell level. We developed SCRIP (**S**ingle-**C**ell Gene **R**egulation Inference using ChIP-seq) (<https://github.com/wanglabtongji/SCRIP>), an integrative method to infer single-cell TR (Transcription Regulator) activity and targets based on the integration of scATAC-seq and a large-scale bulk TR ChIP-seq reference. In comparison to other motif-based approaches, our method performed better when assessing TR binding activity. It also makes it possible to identify TR target genes and build GRNs (Gene Regulation Networks) at the single-cell level using a regulatory potential model. We demonstrate the usefulness of SCRIP in lineage tracing, precise cell-type clustering, and inferring

cell-type-specific GRNs in various biological systems.

- Single-cell RNA-seq technologies made it possible to characterize the immune system heterogeneity in cancers, but also presented computational problems when it came to integrating and using the enormous published datasets to inform immunotherapy. In this work, we present the Tumor Immune Single Cell Hub (TISCH; <http://tisch.comp-genomics.org>), a large curated resource that combines the single-cell transcriptome profiles of approximately 2 million cells from 76 high-quality tumor datasets spanning 27 cancer types. To enable systematic comparison between various cell types, patients, tissue origins, treatment and response groups, and even various cancer-types, TISCH offers interactive gene expression visualization across multiple datasets at the single-cell level or cluster level.

- **Large-scale Genomic Data Mining and Integration Analysis**

Advisors: Dr. Xiaole Shirley Liu & Dr. Clifford Meyer

- We created a parser that can automatically download the raw sequencing data of the chromatin occupancy experiments from the NCBI's GEO database and extract the metadata. The experiment type, tissue type, cell population, and other details of the experiments were annotated and labeled by the parser using the keywords-match strategy.
- After obtaining the raw sequence data, we developed a pipeline, CHIPS (**CH**romatin enrichment **ProceS**sor) (<https://github.com/liulab-dfci/CHIPS>), that can process these epigenome data to readable results with a single command line. It combines state-of-art tools for cistrome NGS data and selects the appropriate analysis parameters. It was developed using the Snakemake framework, which makes it simple to deploy on cloud or cluster servers.
- We developed and maintained the Cistrome Data Browser (<http://cistrome.org/db/#/>). It hosts as many public cistrome sequence datasets, and it provides a user-friendly interface to access the data quality and results of downstream analysis.
- I explored the reads-keeping strategy of ATAC-seq that is used to capture the chromatin accessibility of the genome. I found that only using the fragments with less than 150 bp length can obtain more peaks from ATAC-seq. It was implemented in the CHIPS pipeline.
- Broad signal histone modifications, such as H3K27me3, are difficult to determine the quality of the experiment. We performed unsupervised clustering to mark the low-quality dataset and explore the features of high-quality data and low-quality data.

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

- This research found that the avian influenza virus outbreak was caused by the recombination of two viruses. The virus's conformation changed as a result of recombination, which made it easier for it to bind to human receptors and resulted in the spread of human infection, particularly in the HA protein. It becomes a more adaptable strain as a result of this change in host tropism, which promotes the formation of dominant strains.
- **Screening of Newcastle Disease Virus Heat-Resistant Vaccine by Homologous Recombination**  
*Undergraduate Innovation Program (Presided)*  
Advisor: Dr. Chengqiang He
  - We obtained the live-attenuated viral vaccines from various sources and grew them in a mouse cell line. With the effectiveness of homologous recombination, we cultivated the new heat-resistant vaccine under controlled heat conditions.
- **The Origin, Adaptive Evolution and Population Dynamics of Infectious Chicken Anemia Virus in China**  
*Undergraduate Innovation Program (Involved)*  
Advisor: Dr. Chengqiang He
  - From the NCBI nucleotide database, we obtained the sequences for the Chicken Anemia Virus (CAV). We then used MEGA to align the sequences and create an evolution tree using the maximum likelihood method. We created the dynamic spread path of CAV using the database's metadata, such as where and when the virus was acquired.
- **Diversity of fish in Nansi Lake**  
*Undergraduate Innovation Program (Involved)*  
Advisors: Dr. Rongshu Fu & Dr. Mingsheng Miao
  - Fish species and sizes were gathered and noted in Nansi Lake in Jining, Shandong province. In the meanwhile, we assessed the water quality in Nansi Lake using aquatic organisms like Rotifera, blue-green algae, and others.

## 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.

- **Dong, X.\***; Tang, K.\*; Xu, Y.; Wei, H.; Han, T.; Wang, C. # (2022) Single-Cell Gene Regulation Network Inference by Large-Scale Data Integration. *bioRxiv* July 6, 2022, p 2022.02.19.481131. <https://doi.org/10.1101/2022.02.19.481131>. (Accepted by NAR)
- 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 (7), 1051-1066.e8. <https://doi.org/10.1016/j.stem.2022.06.001>.
- 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

- 2021 Poster Speaking, Human Cell Atlas Asia 2021 Meeting, 2021, Online
- 2021 Poster Speaking, 10<sup>th</sup> National Conference on Bioinformatics and Systems Biology, 2021, Chengdu, China
- 2019 Attendance, 14<sup>th</sup> 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, 14<sup>th</sup> International Bioinformatics Workshop, 2019, Beijing, China

## AWARDS AND DISTINCTIONS

- 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 National Encouragement Scholarship
- 2017 The First Prize of The Fifteenth Challenge Cup Competition in Shandong, Shandong Communist Youth League
- 2017 Advanced Individual of Extracurricular Science and Technology Innovation Activities, Shandong Normal University
- 2017 Excellent Student Cadre in Shandong, Shandong Provincial Education Department
- 2016 The Grand Prize of Shandong College Students' Biochemistry Experiment Skills Competition, Shandong Association of Science & Technology
- 2016 National Encouragement Scholarship
- 2016 Advanced Individual of External Publicity, Shandong Normal University
- 2016 Alumni Scholarship, School of Life Sciences of Shandong Normal University
- 2015 Outstanding Cadres of The Students' Union, Shandong Normal University
- 2015 National Encouragement Scholarship
- 2015 Advanced Individual of Campus Cultural Activities, Shandong Normal University
- 2014 Advanced Individual of Freshman Corps in Military Training, Shandong Normal University

## LEADERSHIP AND PUBLIC SERVICE

- 2022**                    **Volunteer of PCR Test for COVID-19**  
Liferiver Bio-Tech Corp., Shanghai, China
- 2020**                    **Census Taker**  
7<sup>th</sup> 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**                **Organizing Committee Chairman**  
30<sup>th</sup> Top Ten Campus Singers Competition of Shandong Normal University
- 2015**                **Organizing Committee Chairman**  
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