Zheng Wei

✉ weiz@tsinghua.edu.cn

📍 Room 3-425, FIT Building, Tsinghua University, Beijing 100084, China

Bioinformatics | Computational Biology | Deep Learning

RESEARCH EXPERIENCE

Tsinghua University — Beijing, China

Assistant Researcher & Postdoctoral Research Fellow (Supervisor: Xiaowo Wang) — Apr 2021 – Present

- Research on deep learning and eXplainable Artificial Intelligence (XAI)
- Focus on AI, algorithms and software for bioinformatics and synthetic biology (DNA/RNA)

Beijing Academy of Artificial Intelligence — Beijing, China

AI Researcher & Principal Investigator — Jun 2022 – Jun 2023

- AI algorithm and software development for cell-type-specific promoter design
- Manage the project "Intelligent Design of Programmable Medicine"
- Recruit & lead a 6-member team to develop a promoter design/screening platform for gene therapy

Stanford University — Stanford, USA

Visiting Ph.D. student (Supervisor: Wing Hung Wong) — Sep 2018 – Sep 2019

- Deep learning and statistics applied to genomics
- Develop XAI algorithm for decoding cis-regulatory DNA grammar

Tsinghua University — Beijing, China

Ph.D. in Control Science and Engineering (Supervisor: Yanda Li and Xiaowo Wang) — Sep 2014 – Apr 2021

- Research Field: Bioinformatics and Pattern Recognition
- Dissertation: Interpretable Deep Learning for Transcriptional Gene Regulation Analysis

Northeastern University — Shenyang, China

B.Eng. in Computer Science and Technology — Sep 2010 – Jun 2014

- Research Field: Optimization Algorithms and Bioinformatics
- Dissertation: A Tool for Cell Similarity Measurement Based on Gene Network Activity

EDUCATION

Tsinghua University, Stanford University — Beijing, China

Ph.D. in Control Science and Engineering, Department of Automation, Tsinghua University — Sep 2014 – Apr 2021

Visiting student, Department of Statistics, Stanford University, USA — Sep 2018 – Sep 2019

Northeastern University — Shenyang, China

*B.Eng. in Computer Science and Technology; **GPA: 91/100, Rank 3/176*** — Sep 2010 – Jun 2014

RECENT FUNDING

Intelligent Design of Programmable Medicine — Beijing Academy of Artificial Intelligence

Principal Investigator — Jun 2022 – Jun 2023

- Total cost: 10,000,000 CNY ($\approx$1,388,888 USD, 1 USD $\approx$ 7.2 CNY)
- AI-based promoter design and screening for gene therapy
- Intelligent cell-type-specific promoter design algorithm and software
- High throughput promoter screening pipeline development
- Massively parallel reporter assays (MPRA)

Intelligent Design and Control of Biomolecular Machines — Natural Science Foundation of China

One of the three main participants — Jul 2022 – Dec 2024

- Total cost: 2,450,000 CNY ($\approx$340,277 USD, 1 USD $\approx$ 7.2 CNY)
- Original Discovery Program Project
- XAI for genomics and bioinformatics
- Synthetic biology/computational biology software development

Publication

Zheng Wei, Kui Hua, Lei Wei, Shining Ma, Rui Jiang, Xuegong Zhang, Yanda Li, Wing H. Wong, and Xiaowo Wang. "NeuronMotif: Deciphering cis-regulatory codes by layer-wise demixing of deep neural networks." *Proceedings of the National Academy of Sciences* 120, no. 15 (2023): e2216698120.

Zheng Wei[+], Wei Zhang[+], Huan Fang, Yanda Li, and Xiaowo Wang. "esATAC: an easy-to-use systematic pipeline for ATAC-seq data analysis." *Bioinformatics* 34, no. 15 (2018): 2664-2665.

Sijie Chen[+], **Zheng Wei**[+], Yang Chen, Kui Hua, Wei Zhang, Changyi Liu, Haoxiang Gao et al. "SIP: an interchangeable pipeline for scRNA-seq data processing." *bioRxiv* (2018): 456772.

Zheng Wei, Tan Zhu, Tianzhang He, and Shixin Liu. "A fast heuristic algorithm for ladle scheduling based on vehicle routing problem with time windows model." *ISIJ International* 54, no. 11 (2014): 2588-2597.

Honglei Liu, **Zheng Wei**, Antonia Dominguez, Yanda Li, Xiaowo Wang, and Lei S. Qi. "CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation." *Bioinformatics* 31, no. 22 (2015): 3676-3678.

Guiying Wu, Xiangyu Li, Wenbo Guo, **Zheng Wei**, Tao Hu, Yiran Shan, and Jin Gu. "JEBIN: analyzing gene co-expressions across multiple datasets by joint network embedding." *Briefings in Bioinformatics* 23, no. 2 (2022): bbab603.

Patents

Xiaowo Wang, **Zheng Wei.** Training and visualization method and system for extracting and cis-regulatory DNA combination patterns by neural network. Patent CN112735514A. 2021. China

Tan Zhu, **Zheng Wei**, Tianzhang He, Shixin Liu. A digital steel plant equipment control simulation system. Patent CN103217909A. 2013. China

Selected Softwares

NeuronMotif | *Code Ocean* | *GitHub*
- NeuronMotif is an algorithm implemented by Python that can convert the model weight of a deep convolutional neural network (CNN) trained by cis-regulatory DNA sequences into motif grammar (motifs and motif syntaxes).
- It focuses on solving the problem of the multi-faceted neuron to interpret deep convolution neurons in CNN.
- NeuronMotif enables the deciphering of cis-regulatory codes from deep convolutional neurons and enhances the utility of CNN in genome interpretation.

esATAC | *Bioconductor* | *GitHub*
- More than 15 thousand times installation.
- esATAC is a Bioconductor R package that provides a framework and complete preset pipeline for quantification and analysis of ATAC-seq Reads.
- Users can process FASTQ files through the end-to-end preset pipeline which produces a pretty HTML report for quality control and preliminary statistical results, or customize workflow starting from any intermediate stages with esATAC functions easily and flexibly.

CRISPR-era | *Website*
- A fast and comprehensive guide RNA design tool for genome editing, repression and activation.
- The website is developed with JAVA.

Rbowtie2 | *Bioconductor* | *GitHub*
- More than 20 thousand times installation.
- Rbowtie2 is a Bioconductor R package that provide R interface for software of bowtie2 and adapterremoval

Selected Awards

Best Student Paper at the Bioinformatics and Intelligent Information Processing Conference of China: Functional genome annotation and analysis based on chromatin openness and deep learning (oral).

Second Prize Scholarship, Tsinghua University: About the top 20% of students with excellent performance.

China Computer Federation Elite Collegiate Award: The China Computer Federation selects the best 100 computer major college students of the year in China.

Student Science Award: Northeastern University selects the top six undergraduate students who obtained the most outstanding achievements in research in the year.

National Scholarship: China's Ministry of Education rewards students with outstanding academic performance.

Chinese College Students Computer Design Competition, The First Prize (Software Development): Works need to be selected through college competitions (about 2/3) and provincial competitions (165 of 568). National competition are paticipated by 319 colleges in China. In the finals, 19 of the 255 teams of software development participating won the first prize. The estimated winning rate of the first prize is $1\% \sim 3\%$

National College Students Energy Conservation Social Practice and Technology Competition, The Second Prize: Energy conservation competition in China. About the top 8% of the 2051 teams.

"Challenge Cup" Liaoning Province Undergraduate Student Academic Science and Technology Works Competition, Special Prize: About the top 10% of the participants.

Outstanding Graduate, Liaoning Province: About the top 3% of students with excellent academic performance.
Excellent Student (three times), Northeastern University: About the top 10% of students with excellent performance.

First or Second Prize Scholarship of seven terms (three times and four times/seven opportunities in total), Northeastern University: About the top 3% or 10% of students with excellent academic performance.

Skills

Programming: Python, R, C/C++, Java, MATLAB, SQL, HTML

Technologies: At least 5 years of high-performance computer cluster administration and maintenance experience

Languages: Chinese (native), English

Bioinformatics: High throughput sequencing data analysis (e.g. ATAC-seq, DNase-seq, ChIP-seq, RNA-seq)