Chao Hou, Ph.D.
E-mail: ch3849@cumc.columbia.edu, bjmuhou@gmail.com
Website: http://chaohou.netlify.app
1130 St Nicholas Avenue, New York, NY 10032

Academic Curriculum Vitae (Update: May 2025)

Research Interests:

- Large language model for biological sequences
- Protein dynamics and multi-conformation
- Functional analysis of genetic mutations in disease
- Biomolecular interactions and subcellular compartmentalization
- Drug designation and effect prediction
- Transcription regulation and cell fate determination

Work Experience:

- Postdoc Research Scientist — Columbia University — Sep 2023 - Present
  - Advisor: Dr. Yufeng Shen

Education:

- Ph.D. in Biomedical Informatics — Peking University — Sep 2020 - Jul 2023
  - Advisor: Dr. Tingting Li
- Bachelor of Medicine and Economics — Peking University — Sep 2015 - Jul 2020

Main Publications:

(# indicates co-first author, * indicates co-corresponding author)

1. **Hou C***, Di L, Zafar A, Shen Y*. *Understanding Protein Language Model Scaling on Mutation Effect Prediction*. BioRxiv. Apr 2025.
2. **Hou C**, Zhao H, Shen Y*. *Learning Biophysical Dynamics with Protein Language Models* (previous title: *SeqDance: A Protein Language Model for Representing Protein Dynamic Properties*). BioRxiv. Apr 2025.
3. **Hou C#**, Wang X#, Xie H#, Chen T, Zhu P, Xu X, You K, Li T*. *PhaSepDB in 2022: annotating phase separation-related proteins with droplet states, co-phase separation partners and other experimental information.* Nucleic Acids Res. Jan 2023.
4. **Hou C**, Li Y, Wang M, Wu H, Li T*. *Systematic prediction of degrons and E3 ubiquitin ligase binding via deep learning*. BMC Biology. Jul 2022.
5. **Hou C**#, Xie H#, Fu Y#, Ma Y#, Li T*. *MloDisDB: a manually curated database of the relations between membraneless organelles and diseases*. Brief Bioinform. Jul 2021.
6. Chen Z#, **Hou C#**, Wang L#, Yu C, Chen T, Shen B, Hou Y, Li P*, Li T*. *Screening membraneless organelle participants with machine-learning models that integrate multimodal features*. Proc Natl Acad Sci U S A. Jun 2022.
7. Zhu P#, **Hou C**#, Liu M, Chen T, Li T*, Wang L*. *Investigating phase separation properties of chromatin-associated proteins using gradient elution of 1,6-hexanediol*. BMC Genomics. Aug 2023.

8. Han P, **Hou C**, Zheng X, Cao L, Shi X, Zhang X, Ye H, Pan H, Liu L, Li T*, Hu F*, Li Z*. *Serum Antigenome Profiling Reveals Diagnostic Models for Rheumatoid Arthritis*. Front Immunol. Apr 2022.

Other Publications:

9. Xu X, Li Y, Chen T, Hou C, Yang L, Zhu P, Zhang Y, Li T. *investigating variant impact on phosphorylation events driving carcinogenesis*. Brief Bioinform. Dec 2023.
10. Yu C, Lang Y, Hou C, Yang E, Ren X, Li T. *Distinctive Network Topology of Phase-Separated Proteins in Human Interactome*. J Mol Biol. Jan 2022.
11. Shi M, You K, Chen T, Hou C, Liang Z, Liu M, Wang J, Wei T, Qin J, Chen Y, Zhang MQ, Li T. *Quantifying the phase separation property of chromatin-associated proteins under physiological conditions using an anti-1,6-hexanediol index*. Genome Biol. Aug 2021.
12. Chen T, Tang G, Li T, Yanghong Z, Hou C, Du Z, Ma L, Li T. *PhaSeDis: A Manually Curated Database of Phase Separation–Disease Associations and Corresponding Small Molecules*. Genomics Proteomics Bioinformatics. Mar 2025.

Oral Presentations

2024.09 *Learning Representation of Protein Dynamic Properties with a Language Model*. Retreat of Department of Systems Biology of Columbia University. PA, US.

2023.07 *The degradation regulation of phase separating proteins*. The student symposium in Fudan international summer school of life science, Shanghai, China.

2023.06 *The degradation regulation of phase separating proteins*. Excellent graduates symposium of Peking University School of Basic Medical Sciences, Beijing, China.

2023.04 *The degradation regulation of phase separating proteins*. Multidisciplinary Conference on New ideas and New Technologies at Peking University, Beijing, China.

2022.12 *Targeting disordered degrons on phase separating proteins*. Silk Road International Symposium for Distinguished Young Scholars at Xi'an Jiaotong University, Virtual meeting.

Poster Presentations

2024.07 *Learning Representation of Protein Dynamic Properties with a Language Model*. ISMB 2024. Montreal, Canada.

2024.11 *Predicting missense mutation effects with biophysics-informed protein language model*. ASHG 2024. Denver, US.

2025.02 *Learning Biophysical Dynamics with Protein Language Models*. BPS 2025. Los Angeles, US.

Honors and Awards

2023 First prize in the student symposium in Fudan international summer school of life science
2023 Peking University Excellent Graduate
2022 Peking University President Scholarship
2021 Peking University Doctoral Innovation Scholarship
2020 Peking University Junior Scholar
2016 National Scholarship

2014 Silver medal in the final of Chinese Physics Olympiad (high school)

Journal Referee

Ebiomedicine, Plos Computational Biology, Protein Science

Teaching

2024.01 – 2024.09: Rotation student from Department of Biomedical Informatics. Mentor for Aziz Zafar. Columbia Unversity.
2024.06 – 2024.09: Program for Mathematical Genomics (PMG) Undergrad Student Summer Program. Mentor for Jason Xie. Columbia Unversity.
2025.01 – present: Rotation student from Department of Biomedical Informatics. Mentor for Di Liu. Columbia Unversity.