Yuxin Shen

Y.Shen-80@sms.ed.ac.uk • YuxinShen233.github.io,• Google Scholar Page

EDUCATION

University of Edinburgh, School of Biological Science, Edinburgh, UK
PhD student in Biological Science (with Internship) Oct 2022 – present
- BBSRC EASTBIO PhD student
- Research Topic: Machine learning for the optimization of cells

Imperial College London, Department of Chemical Engineering, London, UK
MSc in Advanced Chemical Engineering
Master of Science with Distinction (Grade: 76.26/100) Oct 2020 – Sep 2021
- Core Modules: Modelling of Biological Systems, Advanced Bioprocess Engineering, Advanced Process Operations, Dynamical Systems in Chemical Engineering, Biochemical Engineering, Advanced Process Optimisation
- Master's thesis: Surrogate methods in model-based design of experiments for parameter precision

Fudan University, Department of Chemistry, Shanghai, China
BSc in Chemistry Sep 2016 – Jul 2020
- Core Modules: Physical Chemistry I, II, III, Organic Chemistry I, II, Inorganic Chemistry, Analytical Chemistry I, II, Introduction to Modern Biological Science, Topics in Organic Chemistry, and related experiments
- Selected for Summer Programme in Chemical Biology at University of Hong Kong
- Undergraduate dissertation: Synthesis and phosphorylated peptides enrichment ability of IMAC materials

SCIENTIFIC RESEARCH EXPERIENCES

PhD Research
School of Biological Science, University of Edinburgh Oct 2022 – present
Supervisor: Prof. Diego Oyarzún and Prof. Grzegorz Kudła

Machine learning for sequence-to-expression mapping
- Integrated biological domain knowledge into feature engineering for machine learning models.
- Worked on both supervised learning (including RF, MLP, CNN, GNN and foundational DNA language models) and unsupervised learning for sequence-to-expression modelling.
- Built a machine learning model for membrane protein expression with high prediction accuracy and explainability.

Active learning for biological sequence optimization
- Created an active learning pipeline for sequence generation and selection for optimal cellular performance.

MSc Research
Department of Chemical Engineering, Imperial College London Nov 2020 – Sep 2021
Supervisor: Prof. Benoit Chachuat

A comparative study of surrogate methods in model-based design of experiments
- Applied different surrogates (Artificial Neural Network, Bayesian Neural Network) to approximate model sensitivities; conducted model-based design of experiments.

RESEARCH OUTPUT

1. Shen, Y., Kudla, G. and Oyarzún, D.A., 2025. Improving the generalization of protein expression models with mechanistic sequence information. *Nucleic Acids Research*, 53(3), gkaf020.
2. Shen, Y., Kudla, G. and Oyarzún, D.A., 2025. Optimization of regulatory DNA with active learning. *Computational and Structural Biotechnology Journal*, 27, pp.4384-4392.
3. Shen, Y., Underhill, J., Mulholland, A.J., Oyarzun, D.A. and Curnow, P., 2025. Effective sequence-to-expression prediction for membrane proteins using machine learning and computational protein design. *bioRxiv*, pp.2025-09.
4. Çubuk, H., Plech, M., Aslanzadeh, V., Zikanova, M., Skopova, V., Kmoch, S., Shen, Y., Marsh, J.A. and Kudla, G., 2025. Mechanistic Modelling of Recessive Disease through Allelic Integration of Variant Effects. *bioRxiv*, pp.2025-08.

WORKING & INTERNSHIP EXPERIENCES

Informatics (Machine learning) Intern
Etcembly, Oxfordshire, UK May 2025 – Jul 2025
- PIPS (Professional Internships for PhD Students) of my PhD
- Developed and implemented a pipeline of Large Language Models (LLMs) + biological domain knowledge for antibody and TCR sequence design.

- Designed sequences with enhanced performance in wet lab validation.

R&D Technologist
Research and Development, Unilever (China) Shanghai Branch Jan 2022 – May 2022
- Worked on product development (formulation development) in laundry products for Homecare sector
- Focused on new functions and sustainable solutions in liquid/concentrated laundry product formulation

Bioprocess Engineering Intern
Pall Corporation, Danaher Life Science Early Career Program Sep 2021 – Dec 2021
- Assisted upstream bioprocess design for gene therapy and monoclonal antibody production
- Designed single-use systems for bioreactors and the filtration processes

Bioanalysis Intern
Industrial Research Institute, Shanghai Fosun Pharmaceutical (Group) Co., Ltd. Dec 2019 – Feb 2020
- Assisted the PK/PD bioanalysis of the APIs for the generic drugs
- Reviewed LC-MS methodology validation reports and bioanalytical reports

HONORS & AWARDS

BBSRC EASTBIO PhD studentship (4-year full scholarship) Jun 2022
Graduate Scholarship of Department of Chemistry (The Third Prize) Jun 2020
Professional Scholarship of Department of Chemistry Oct 2018 & Oct 2017
Outstanding Student Scholarship of Fudan University (The Third Prize) Oct 2019 & Oct 2018 & Oct 2017
First Prize in Chinese Chemistry Olympiad in Shanghai Oct 2015

OTHER EXPERIENCES & PROJECTS

Turing Institute Interest Group seminar series: Co-organizer Sep 2024 – present
- Co-organize the Data-Centric Biological Design & Engineering Interest Group
- Invite speakers and organize seminars
- Update the seminar website, advertise the seminar series

BBSRC EASTBIO DTP Student Representative Oct 2022 – Sep 2024
EASTBIO student representative for College of Science and Engineering, University of Edinburgh
- Organized panel and poster sessions in Annual EASTBIO Symposiums
- Worked on Q&A information sessions for new students and potential applicants

EASTBIO lead student representative on industrial engagement committee
- Worked on maintaining the network between industrial partners and PhD students

PROGRAMMING SKILLS

Programming: Python, Julia, R, Octave
- Python Packages for Scientific Computing and Machine Learning: NumPy, Pandas, SciPy, NetworkX, Matplotlib, Seaborn, Scikit-Learn, PyTorch, PyTorch Geometric, Tensorflow, JAX
- Julia Packages: DifferentialEquations, DiffEqFlux, Optimization, Catalyst
- Shell Command line operations, Git, basic software testing and packaging

Deep Learning for Healthcare Specialization Certificate on Coursera
Deep Learning Specialization Certificate on Coursera
Machine Learning Specialization Certificate on Coursera

ATTENDED CONFERENCES & WORKSHOPS

Bioprocessing Summit Europe 2025: Speaker Mar 2025
Synthetic Biology UK 2024: Oral Communication Nov 2024
Protein & Antibody Engineering Summit (PEGS) Europe 2024: Poster presentation Nov 2024
Synthetic Biology for Health and Sustainability: Short talk Oct 2024
EASTBIO Annual Symposium 2024: Student representative and Co-Organizer Jun 2024
ChemEngDayUK 2024: Flash talk and poster presentation Apr 2024
EMBL Symposium AI and biology: Flash talk and poster presentation Mar 2024
BBSRC Inter-DTP Module in Software Engineering Feb 2024
Synthetic Biology UK 2023: Poster presentation Nov 2023
ELLIS Summer School on Machine Learning for Healthcare and Biology Jun 2023
EASTBIO Annual Symposium 2023: Student representative and Co-Organizer Jun 2023
AI, Engineering Biology & Beyond - Turing Institute Workshop: Volunteer in logistics Mar 2023
Host of EASTBIO Thematic Meeting “Computational tools for integrated Omics” Feb 2023