QIJIE HE

qijiehe01@gmail.com | https://github.com/Nadine-sudo

EDUCATION

Sun Yat-sen University, Shenzhen, China

Degree: Master of Science in Bioinformatics

Time: Sep. 2022 - June. 2025

Relevant courses: Oncogenomics, Bioinformatics, Big Data Analysis

China University of Geosciences, Wuhan, China

Degree: Bachelor of Science in Biological science

Time: Sep. 2017 - Jun. 2021

Relevant courses: Cell Biology, C Language Program Design, Advanced Mathematics, R Language and Biostatistics, Probability and Mathematics Statistics, Biochemistry, Genetics, Molecular Biology

RESEARCH EXPERIENCE

Project: Mutational signatures of N-nitrosamines in human cancers

Time: Jan. 2023 – present

Key Contributions:

- Independently established 2 in vitro cell models and 1 in vivo rat tumor model to characterize N-nitrosamine-induced mutation patterns.
- Processed whole-genome sequencing samples using GATK variant calling and integrated experimental data with tumor patient data from public databases, identified N-nitrosamine-specific mutations induced by mutagenic exposure, and verified the presence of these novel mutations in patient cohorts.
- Developed a modified nonnegative matrix factorization (NMF) pipeline to extract novel N-nitrosamine-related mutation signatures from sequencing data.

Presentation: Cell Symposia: Hallmarks of cancer, Guangzhou, China, Nov. 2024, **Qijie He**, Runtian Yao, Mini Huang. "*Mutational signatures of N-nitrosamines in human cancer*" (poster)

Project: MethylMSI: Prediction of microsatellite instability based on DNA methylation profile and SVM model

Time: Dec. 2024 – Jun. 2025

Key Contributions:

- Developed MethylMSI — an SVM-based tool for microsatellite instability (MSI) classification—using TCGA DNA methylation data, and optimized its user-friendliness; open-sourced it at https://github.com/Huanglab-ai/MethylMSI.
- Compared MethylMSI's classification performance with existing methods across multiple TCGA cohorts, demonstrating its superior performance relative to other tools.
- Contributed to writing key sections of the study, including detailing MSI's role in cancer immunotherapy, analyzing the SVM model's core features, and affirming SVM's potential as the optimal MSI prediction model.

Paper: Chenzhi Yan#, **Qijie He**#, Runtian Yao, Jiawen Wu, Xing-Ding Zhang, Lin Qi*, Mi Ni Huang*. *MethylMSI: Prediction of microsatellite instability based on DNA methylation profile and SVM model*. Computational Biology and Chemistry, Submitted, 2025.

Project: Multi-tissue gene expression dysregulation in the progression of Pompe disease mouse models
Time: Nov. 2022 – Jun. 2024
Key Contributions:
- Applied time-series cluster analysis on differentially expressed genes using the Mfuzz package's soft clustering algorithm, identifying dynamic expression patterns across disease stages.
- Conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses.
- Performed alternative splicing analysis with the IsoformSwitchAnalyzeR package, unraveling isoform-level complexity in gene expression regulation during disease progression.

Paper: Jialiang Luo#, **Qijie He**#, Runtian Yao, Zifu Ding, Jiawen Wu, Suning Ping*, Mi Ni Huang*. *Multi-tissue gene expression dysregulation in the progression of Pompe disease mouse models*. Molecular Genetics and Metabolism, Submitted, 2025.

TECHNICAL SKILLS

Computational & Programming Skills:
Linux: Proficient in shell scripting.
Python: Skilled in data analysis (Pandas, NumPy) and visualization (Matplotlib, Seaborn), capable of writing basic web crawlers (via requests, BeautifulSoup libraries) to retrieve public data.
R: Proficient in various bioinformatics-related packages, with a focus on transcriptome analysis (clusterProfiler, DESeq2) and visualization (ggplot2).
C: Equipped with foundational knowledge from core courses; able to independently write simple programs for basic data processing tasks.
Machine Learning Workflows & Frameworks: Familiar with ML workflows, proficient in PyTorch for basic model building and Scikit-learn for classification/regression tasks, also has basic knowledge and hands-on experience in natural language processing-related operations.

Wet-Laboratory Techniques:
Animal Experimentation: Experienced in establishing mouse/rat tumor models and conducting animal administration (oral gavage, intraperitoneal injection, biological sample collection).
Cell Culture: Skilled in maintaining adherent cell lines with standard operations including cell passage, cryopreservation, and thawing; capable of cell viability assessment.
Molecular Techniques: Proficient in preparing sequencing libraries (NGS) for Illumina platform, skilled in basic molecular operations including DNA extraction (from cells/tissues), RNA extraction (TRIzol method), agarose gel electrophoresis (nucleic acid quality verification), and PCR amplification.