

# Mengting Li

PHD STUDENT IN BIOINFORMATICS

Room 3-08, The Hong Kong Jockey Club Building For Interdisciplinary Research, 5 Sassoon Rd, Sandy Bay, Hong Kong

she/her | +852 5968 9107 | mtlee1001@connect.hku.hk | limengting.com | limengting | limengting | limengting

## Education

### The University of Hong Kong (HKU)

PHD IN BIOINFORMATICS

Hong Kong, China

Oct 2025 - Present

### University of Chinese Academy of Sciences (UCAS)

MASTER OF BIOLOGICAL AND PHARMACEUTICAL ENGINEERING

Beijing, China

Sep 2020 - Jun 2023

### Xuzhou Medical University (XZMU)

BACHELOR OF BIOINFORMATICS

Xuzhou, China

Sep 2016 - Jun 2020

## Research Experience

### Research Assistant (Jul 2023 - Sep 2025)

#### The University of Hong Kong | Hong Kong, China (Supervisor: Prof. Leo Poon)

##### Project: An interferon-stimulated long non-coding RNA USP30-AS1 as an immune modulator in influenza A virus infection

- Developed a reproducible bioinformatics pipeline that successfully processed and analyzed 5 RNA-seq datasets and 4 microarray datasets from public repositories (SRA, GEO).
- Identified 1,715 differentially expressed lncRNAs across multiple post-infection time points by implementing a custom differential expression analysis workflow (using STAR, featureCounts, and DESeq2).
- Revealed key immune pathways using g:Profiler for functional enrichment analysis and EnrichmentMap in Cytoscape for network visualization.

### Master's Student Researcher (Sep 2020 - Jun 2023)

#### BGI-Research & University of Chinese Academy of Sciences | Shenzhen, China (Supervisor: Prof. Kui Wu)

##### Project: Multi-omics sequencing of gastroesophageal junction adenocarcinoma reveals prognosis-relevant key factors and a novel immunogenomic classification

- Led a multi-omics analysis of 92 gastroesophageal junction adenocarcinoma (GEJAC) patients, integrating whole-genomic, transcriptomic, and T-cell receptor repertoires data.
- Developed a comprehensive bioinformatics pipeline for processing WGS data to perform somatic variant calling and driver gene identification. Further utilized these results for downstream analyses including mutational signature analysis (SigProfilerExtraction), neoantigen prediction (NetMHC, NetMHCpan), and tumor evolution (PhylogicNDT).
- Discovered 30 mutated driver genes, 7 novel genomic signatures, 3 copy-number variations, and 2 V-J gene usages previously unidentified and linked to prognosis.
- Developed a novel immunogenomic classification system that stratifies patients into three immune-related subtypes (immune-inflamed, intermediate, and deserted) with discrete profiles of genomic signatures, immune status, and clinical outcomes.

## Skills

### Programming & Scripting Languages

- R:** (Proficient) Data manipulation (Tidyverse, data.table), statistical modeling, bioinformatics (Bioconductor), visualization (ggplot2).

- **Python:** (Proficient) Scientific computing (Pandas, NumPy), machine learning (PyTorch, Scikit-learn), scripting and workflow integration.
- **Bash/Shell Scripting:** Proficient in automating tasks, managing data, and working in a Linux/HPC environment.

## Bioinformatics & Genomics

- **Omics Analysis & Pipeline Development:** Proficient in building and deploying automated analysis pipelines for diverse genomic data modalities, including Whole Genome Sequencing (WGS), Bulk RNA-seq, single-cell RNA-seq (scRNA-seq), and Spatial Transcriptomics.
- **Phylogenetics & Evolutionary Analysis:** Experience in Bayesian phylogenetic inference (BEAST), maximum likelihood methods (e.g., IQ-TREE), and population genetics simulation (SLiM).
- **Sequence Analysis:** Viral lineage assignment (Pangolin), sequence alignment (BWA, Bowtie2), variant calling & manipulation (VCFtools, BCFtools), and general sequence analysis (Samtools, BLAST).
- **Genomic Data Formats:** Deep familiarity with standard formats including VCF, BAM/SAM, FASTQ, FASTA, and GFF/GTF.

## Developer Tools & Workflow Management

- **Version Control:** Git / GitHub for collaborative and reproducible research.
- **Environment Management:** Conda for managing software dependencies and creating reproducible environments.
- **High-Performance Computing (HPC):** Experience with job scheduling systems (e.g., Slurm, SGE) and working in a cluster environment.
- **Workflow Languages:** Familiarity with the principles of workflow management using Snakemake / Nextflow.

## Deep Learning

- **Machine Learning Foundations:** Strong foundation in Regression, Classification, Clustering, DNNs, and CNNs.
- **Advanced Sequence Models:** Understanding of advanced architectures including RNNs, LSTMs, and Transformers.
- **Technical Implementation:** Hands-on project experience applying Scikit-learn, PyTorch, and TensorFlow to bioinformatics tasks.

## Publications

---

1. Ma, Z., Li, M., Li, F., Wu, K., Wu, X., Luo, T., Gao, N., Luo, H., Sui, Z., Yu, Z., Jiang, H., Shang, X., Chen, C., Yue, J., Meng, F., Duan, X., & Xu, B. (2025). Multi-omics sequencing of gastroesophageal junction adenocarcinoma reveals prognosis-relevant key factors and a novel immunogenomic classification. *Gastric Cancer*, 28(3), 344–357. <https://doi.org/10.1007/s10120-025-01585-y>
2. Cao, Y., Chin, A. W. H., Gu, H., Li, M., Gu, Y., Lau, S. P. N., Hui, K. P. Y., Chan, M. C. W., & Poon, L. L. M. (2025). An interferon-stimulated long non-coding RNA USP30-AS1 as an immune modulator in influenza A virus infection. *PLoS Pathog*, 21(1), e1012854. <https://doi.org/10.1371/journal.ppat.1012854>
3. Cao, Y., Chin, A. W. H., Gu, H., Li, M., & Poon, L. L. M. (2025). Long non-coding RNA BCAR4 is required for efficient influenza A virus replication. *Microbiology*. <https://doi.org/10.1101/2025.02.06.636885>
4. Zhang, J., & Li, M. (2020). Bioinformatics analysis and design based on r-studio and databases. *Modern Information Technology*, 4(4):76–79. <https://qikan.cqvip.com/Qikan/Article/Detail?id=7101730810>