

Qiyang Hong

Institute of Basic Medical Sciences, Peking Union Medical College & Tsinghua University School of Medicine

Beijing, China

Email: qiyanghong2020@gmail.com

GitHub: <https://github.com/qiyanghong2020>

Google Scholar: <https://scholar.google.com.hk/citations?user=1PCtyx8AAAAJ&hl=zh-CN>

Personal site: <https://qiyanghong2020.github.io/>

EDUCATION

- **Ph.D. in Biomedical Engineering (Medical AI)**, Peking Union Medical College & Tsinghua University School of Medicine, Beijing, China (Sept 2023–Present)
 - Advisor:** Prof. Erping Long (Researcher; National Overseas Young Talent)
 - Core Expertise:** Foundation models, LLMs, multimodal learning, temporal modeling; deep phenotyping, clinical AI; model interpretability; high-performance computing (4×A6000, 2×H100)
- **M.S. in Biochemistry and Molecular Biology**, Xiamen University, Xiamen, China (Sept 2014–July 2017)
 - Advisor:** Prof. Ning-Shao Xia (Academician, Chinese Academy of Engineering)
 - Specialization:** Transcriptomics, viral immunology, structural biology, computational molecular analysis
- **B.Eng. in Bioengineering**, Jimei University, Xiamen, China (Sept 2010–July 2014)

TECHNICAL SKILLS

- **Artificial Intelligence & Machine Learning:**
Neural networks, Transformer architectures, foundation models, LLMs, multimodal learning, temporal modeling, representation learning, model interpretability.
- **Model Training & High-Performance Computing:**
PyTorch, CUDA, distributed training, GPU cluster management (Linux HPC, A6000 nodes), job scheduling, system optimization, large-scale data preprocessing, performance profiling.
- **Bioinformatics & Genomic Analysis:**
WES/WGS pipelines, RNA-seq, SNP/Indel/CNV detection, variant annotation and ACMG classification, immunogenomics, neoantigen prediction, clinical diagnostics workflow development.
- **Programming & Data Engineering:**
Python, NumPy, pandas, scikit-learn; workflow automation; reproducible pipelines; Django-based web applications.
- **Data Visualization & Network Analysis:**
matplotlib, seaborn, igraph, networkx; interactive visualization for clinical and genomic datasets.

CERTIFICATIONS & LICENSES

- **Health Professional Qualification (Clinical Laboratory / Medical Testing)**
National Health Authority of China — 2012
(National-level professional technical qualification)
- **Clinical PCR Laboratory Technician Certification**
Fujian Provincial Clinical Laboratory Center — Sept 2020
(Authorized training for clinical gene amplification laboratories)
- **Bioinformatics Engineer Certification**
ICT Support / ICTTT — Jan 2015

PROFESSIONAL EXPERIENCE

- **Group Leader, Molecular Diagnostics**
Affiliated Hospital of Putian University, Prenatal Diagnosis Center
Aug 2020 – Jul 2023
- **Senior Bioinformatics Engineer (Part-time)**
Jinfeng Biotechnology Co., Ltd., Data & Information Department
Nov 2020 – Jan 2022
- **Bioinformatics Supervisor**
Berry Genomics Co., Ltd.
Jul 2017 – Aug 2020

PUBLICATIONS

*Equal contributor, #Corresponding author

In revision/ review/ preparation:

1. [Hong Q*](#), [Xu S*](#), [Zhang Q*](#), ..., [Long E#](#). [Educator–LLM collaboration strategies for discharge education in hematopoietic stem cell transplantation: a randomized controlled trial of real-time prompting.](#) *NEJM AI* (Under Revision)
2. [Ni X*#](#), [Chen T*](#), [Hai E*](#), [Huang Z](#), [Li Y](#), [Sun X](#), [Hong Q](#), [Long E#](#). [A mediation-informed causal pathway powers a large language model for individualized health advisory.](#) *Nature Aging* (Under Review)
3. [Wu F*](#), [Hong Q*](#), [Pang J*](#), ..., [Long E#](#), [Ran P#](#), [Yi E#](#). [Cough, sputum, and CD164: Novel risk factors/biomarkers for COPD and lung function decline.](#) *Respiratory Research* (Submitted)
4. [Pang J*](#), [Ren X*](#), [Hong Q*](#), ..., [Long E#](#). [Proteo-metabolomic insights into the progression of COPD and lung function decline.](#) *Respiratory Research* (Under Review)
5. [Pang J*](#), [Wang J*](#), [Wu S](#), [Hong Q](#), ..., [Long E#](#). [GWAS meta-analyses with a new Chinese population expand the genetic architecture and improve ancestry-specific genetic risk prediction of chronic obstructive pulmonary disease.](#) *Genomics Proteomics and Bioinformatics* (In press)

Published/ accepted:

1. [Hong Q*](#), Wang C*, ..., Long E#. A foundational model encodes deep phenotyping data and enables diverse downstream applications. *npj Digital Medicine* 2026. DOI: 10.1038/s41746-026-02736-w.
2. Pang J*, Wang J*, Wu S, [Hong Q](#), ..., Long E#. GWAS meta-analyses with a new Chinese population expand the genetic architecture and improve ancestry-specific genetic risk prediction of chronic obstructive pulmonary disease. *Genomics Proteomics and Bioinformatics* 2026 (In press)
3. Sun X*, [Hong Q*](#), ..., Wan P#, Long E#. Model confrontation and collaboration: a debate intelligence framework for enhancing medical reasoning in large language models. *Cell Reports Medicine* 2025 (IF 10.6)
4. Ran P*, Yi E*, Cui J*, Wang H*, [Hong Q*](#), ..., Long E#. Multi-Omics Mendelian Randomization Identifies SERPING1 as a COPD Modulator. *Signal Transduction and Targeted Therapy* 2025 (IF 52.7)
5. Long Y*, Ni X[#], Chen T*, Huang X, [Hong Q](#), ..., Zhang T#, Long E#. Genetic determinants of gene expression noise and its role in complex trait variation. *Cell Reports* 2025
6. Wu W*, Zhou M*, Chen J, Wang J, [Hong Q](#), ..., Pang J#, Long E#, Wang J#. Deep multi-omics profiling reveals three distinct molecular subtypes of chronic obstructive pulmonary disease in a unique biomass-exposed Chinese population. *Med* 2025
7. Lin K*, [Hong Q*](#), ..., Chen M#. Cervical HPV infection and related diseases among 149,559 women in Fujian: an epidemiological study from 2018 to 2023. *Frontiers in Microbiology* 2024
8. Huang Y*, Li Z*, [Hong Q](#), ..., Yu H#, Li S#, Xia N#. A stepwise docking molecular dynamics approach for simulating antibody recognition with substantial conformational changes. *Computational and Structural Biotechnology Journal* 2022
9. Xin L*, Yu H, [Hong Q](#), ..., Xia N#. Identification of strategic residues at the interface of antigen–antibody interactions by in silico mutagenesis. *Interdisciplinary Sciences: Computational Life Sciences* 2018
10. Xu L*, Zheng Q, Li S, He M, Wu Y, Li Y, Zhu R, Yu H, [Hong Q](#), ..., Xia N#. Atomic structures of Coxsackievirus A6 and its complex with a neutralizing antibody. *Nature Communications* 2017
11. Li Z*, Yan X*, Yu H*, Wang D, Song S, Li Y, He M, [Hong Q](#), ..., Baker T#, Li S#, Xia N#. The C-terminal arm of the human papillomavirus major capsid protein is immunogenic and involved in virus-host interaction. *Structure* 2016
12. Zhang X*, Wang K*, Lin Q*, Zheng M, Li Q, Li T, [Hong Q](#), ..., Yu H#, Li S#, Xia N#. A shared N-terminal hydrophobic tail for the formation of nanoparticulates. *Nanomedicine* 2016

PROFESSIONAL ACTIVITIES

Reviewer (ad-hoc): Nature Medicine, Nature Biomedical Engineering, Frontiers in Aging Neuroscience

RESEARCH PROJECTS & TECHNOLOGY DEVELOPMENT

● **Temporal Foundation Models for COPD Progression**

National Science and Technology Major Project (Youth Scientist Program), 2025–2028

Developing longitudinal multi-omics integration and temporal neural networks to decode COPD progression, identify molecular signatures, and support early risk prediction.

- **Biomedical Foundation Model (UK Biobank)**

Built a Transformer-based foundation model trained on deep phenotyping data from >500,000 participants, enabling disease prediction, multimorbidity analysis, and patient stratification across 289 conditions.

- **Automated Clinical WES Analysis Pipeline**

Developed an end-to-end workflow from FASTQ to SNP/Indel/CNV detection, annotation, ACMG classification, and automated reporting; deployed in clinical settings with version-controlled pipelines.

- **Neoantigen Prediction and Immunogenomics Pipeline**

Integrated WES and RNA-seq pipelines with NetMHC/MHCflurry to identify candidate neoantigens for immunotherapy and tumor profiling.

- **WES Clinical Interpretation and Visualization Platform**

Implemented a Django-based interactive web platform enabling clinicians to visualize variants, coverage, CNVs, and Sanger traces, supporting clinical genetic diagnosis.

- **Consumer Genomics Analysis Framework**

Contributed to the development and maintenance of large-scale production pipelines for SNP genotyping and ancestry/trait reporting in collaboration with consumer genomics companies.

- **Inherited Disease and Metabolic Panel Pipelines**

Designed and optimized pipelines for thalassemia, SMA, mitochondrial disorders, and mtDNA variant detection to support clinical molecular diagnostics.

- **Custom WES Probe Design Optimization**

Improved exome capture efficiency by designing probe sets enriched for pathogenic ClinVar regions, enhancing diagnostic yield.

- **High-throughput Transcriptomics Analysis**

Performed differential expression and KEGG/GO enrichment analysis for bacterial and human datasets under diverse physiological conditions.