

# XIAOTONG YAO

+1(917)833-5729 • xiy2006@med.cornell.edu • xiaotong.yao23@gmail.com

Github, Twitter, LinkedIn: xtYao • ORCID: 0000-0002-5140-6639

1300 York Ave • New York, NY 10065

## EDUCATION

### Weill Cornell Medicine

Jul 2015- expected Apr 2021

Ph.D. in Computational Biology & Medicine

Thesis (supervisor: Marcin Imielinski): *Illuminating rearranged cancer genome structures through genome graphs*

### New York University

Aug 2013-May 2015

M.S. in Bioinformatics & Systems Biology (GPA: 3.9/4)

Thesis (supervisor: Christine Vogel): *iSUMO - integrative prediction of functionally relevant SUMOylated proteins*

### Zhejiang University

Aug 2009-May 2013

B.S. in Biotechnology (GPA: 3.4/4)

## EXPERIENCE

### Weill Cornell Medicine

Jul 2015 - Present

Graduate Research Assistant

New York, NY

- Discovered three new patterns of complex structural variations from thousands of cancer genomes
- Designed *Junction Balance Analysis* to reconstruct genome graphs
- Conceptualized and developed *gGnome* for genome graph data structure
- Contributed to large cancer sequencing consortiums including TCGA and PCAWG

### New York University

December 2013 - May 2015

Graduate Research Assistant

New York, NY

- Trained predictive models for post-translational modification from protein function databases

### New York University

Feb 2014 - May 2015

Teaching Adjunct

New York, NY

- Taught R statistical programming for Introduction to Biostatistics
- Tutored techniques for Biological Databases and Data Mining

### 3E Bioenergy

Jun 2014 - May 2015

Bioinformatics Intern

New Brunswick, NJ

- Compared crop genomes for candidate genes linked to drought resistance in sweet sorghum

## SELECTED PUBLICATIONS

Hadi K, **Yao X**, Behr JM, ..., Imielinski M. *Distinct Classes of Complex Structural Variation Uncovered across Thousands of Cancer Genome Graphs*. Cell. 2020;183: 197–210.e32.

(In review) Dewhurst SM, **Yao X**, ..., de Lange T, Imielinski M. *Structural variant evolution after telomere crisis*. BioRxiv. 2020. p. 2020.09.29.318436. doi:10.1101/2020.09.29.318436

(In review) Carrot-Zhang J, **Yao X**, Devarakonda S, et al. *Whole-genome characterization of lung adenocarcinomas lacking alterations in RTK/RAS/RAF/MAPK pathway*. Cancer Res. 2020;80: 5895–5895.

## SKILLS

### Biology

Cancer Genomics, Computational Biology, Systems Biology

### Sequencing Informatics

Illumina, Oxford Nanopore WGS

### Statistics & Machine Learning

Generalized Linear Models, Random Forests, Regularized Regression

### Computer Languages

R, Shell, Python, Java

### Optimization

CPLEX, Gurobi

### Data Visualization

ggplot, shiny

### Scientific Communication

LaTeX, Adobe Illustrator, (R)markdown

### Databases

SQLite, MySQL

### Other Tools

Git, Emacs, Docker, Nextflow, Hugo