Xiaoran Tong

1023 Saffron Loop, Durham, NC ♦ Phone: 5172201229 ♦ Email: xiaoran.tong.cn@gmail.com

EXPERIENCE

09/2020 – present National Institutes of Health Durham, NC

Post-doc Research Fellow, Biostatistics, Epidemiology

- variance loci analysis (VLA) to select genes highly interactive with environmental exposure[1] (methodology development).
- catalog gene-environment interaction candidate for over 3,000 UK Biobank phenotypes through VLA[2] (high performance computation, data management).
- genetic risk prediction for insomnia in youth of diverse ancestry [3] (mentorship training).

05/2014 – 05/2020 MICHIGAN STATE UNIVERSITY East Lansing, MI

Graduate Research Assistant, Statistical Genetics, Epidemiology

- (Ph.D.) Develop a neural network method to process cortex image in the search of Alzheimer's Disease gene[4].
- (Ph.D.) Conduct gene-level GWAS with NGS and gene expression data to detect hypertension causal genes[5].
- (Ph.D.) Develop a kernel method handling Bio-Bank sized genotype to improve disease prediction [6], [7].
- Develop Python tool "hpcwp" to aid cluster computation under SLURM environment[8].
- Develop R-package "plinkFile" to load genotype count data and kinship data matrices[9].
- Repurpose the Netflix film recommender for imputing missing genotype counts.
- Analyze Michigan Twin Registry for anti-social behaviors[10, p.] (high-dim count data analysis);
- Provide support in statistics, data preparation, computation, and program troubleshooting [11]–[17].

06/2010 – 01/2014 ZHEJIANG UNIVERSITY Hangzhou, Zhejiang, China

Research Assistant, Bioinformatics

- Develop biofabric visualization for multi-way genetic interaction using Java and Java DB [18], [19].
- bioinformatics training.
- practice in software engineering (version control, unit test, and documentation).

09/2007 – 06/2008 YUNNAN MINZU UNIVERSITY Kunming, Yunnan, China

Instructor, Computer Science

- Instructor for C/C++ programming, data structure, and computer algorithm.

EDUCATION

2014 – 2020 MICHIGAN STATE UNIVERSITY East Lansing, Michigan, USA

Ph.D. in Epidemiology and Biostatistics

- research focus: machine learning methods in human genetics

2010 – 2014 ZHEJIANG UNIVERSITY Hangzhou, Zhejiang, China

Ph.D. Candidate in Bioinformatics

- study bioinformatics, genetics, and molecular biology

2000 – 2009 KUNMING UNIVERSITY of SCI & TECH Kunming, Yunnan, China

Master and Bachelor, Computer Science

SKILLS

- Proficient in program with R, Python, Linux shell, and C++.
- Practitioner of machine learning and classical statistical methodologies.
- Familiar of public databases (e.g., ENCODE, ENSEMBL, UK Biobank) and GWAS pipeline.
- Proficient in software for statistical genetics (e.g., PLINK, Bcftools, GCTA, Mimimac, etc.);
- Familiar with database query by SQL.
- Skills in communication, documentation, and presentation.

Published Work and Software Repositories

[1] T. Xiaoran and M. Alison, "Identifying GxE Variants with Variance Locus Analysis," *NIH/NIEHS*. https://genelist.niehs.nih.gov/vla/about (accessed Apr. 21, 2022).

Xiaoran Tong

1023 Saffron Loop, Durham, NC ♦ Phone: 5172201229 ♦ Email: xiaoran.tong.cn@gmail.com

[2] X. Tong and A. Motsinger-Reif, "The Variance Locus Analysis Catalog: An open database of Genome-wide Gene-Environment Mapping." NIEHS, 2020. [Online]. Available: https://edelgene.niehs.nih.gov/vla/about

[3] T. Ma, H. Chen, Q. Lu, and X. Tong, "Polygenic Risk for Insomnia in Adolescents of Diverse Ancestry," *Frontiers in genetics*, vol. 12, 2021.

[4] X. Tong and Q. Lu, "A non-parametric method for joint association analysis of sequencing and Imaging data," in *GENETIC EPIDEMIOLOGY*, 2015, vol. 39, pp. 586–586.

[5] X. Tong, C. Wei, and Q. Lu, "Genome-wide joint analysis of single-nucleotide variant sets and gene expression for hypertension and related phenotypes," in *BMC proceedings*, 2016, vol. 10, p. 36.

[6] X. Shen, X. Tong, and Q. Lu, "A Kernel-Based Neural Network for High-dimensional Genetic Risk Prediction Analysis," *arXiv:2101.11807 [stat]*, Jan. 2021, Accessed: Apr. 19, 2022. [Online]. Available: http://arxiv.org/abs/2101.11807

[7] M. Tang *et al.*, "Fast heritability estimation based on MINQUE and batch training," *Briefings in Bioinformatics*, p. bbac115, Apr. 2022, doi: 10.1093/bib/bbac115.

[8] X. Tong, *A HPCC wrapper to create and manage paralell jobs in a PBS/SLURM computation environment*. 2022. Accessed: Apr. 21, 2022. [Online]. Available: https://github.com/xiaoran831213/hpcwp

[9] X. Tong, *R-package: (Github) plinkFile*. 2022. Accessed: Mar. 31, 2022. [Online]. Available: https://github.com/xiaoran831213/plinkFile

[10] J. J. Tielbeek *et al.*, "Genome-Wide Association Studies of a Broad Spectrum of Antisocial Behavior," *JAMA Psychiatry*, vol. 74, no. 12, pp. 1242–1250, Dec. 2017, doi: 10.1001/jamapsychiatry.2017.3069.

[11] O. A. Vsevolozhskaya, D. V. Zaykin, D. A. Barondess, X. Tong, S. Jadhav, and Q. Lu, "Uncovering local trends in genetic effects of multiple phenotypes via functional linear models," *Genetic epidemiology*, vol. 40, no. 3, pp. 210–221, 2016.

[12] S. Jadhav, X. Tong, and Q. Lu, "A functional U-statistic method for association analysis of sequencing data," *Genetic epidemiology*, vol. 41, no. 7, pp. 636–643, 2017.

[13] S. Jadhav, O. A. Vsevolozhskaya, X. Tong, and Q. Lu, "The impact of genetic structure on sequencing analysis," in *BMC proceedings*, 2016, vol. 10, p. 60.

[14] C. Wu *et al.*, "Novel strategy for disease risk prediction incorporating predicted gene expression and DNA methylation data: a multi-phased study of prostate cancer," *Cancer Communications*, vol. 41, no. 12, pp. 1387–1397, 2021, doi: 10.1002/cac2.12205.

[15] S. A. Burt, K. Klump, Q. Lu, and X. Tong, "A straightforward family design for examining the contribution of mitochondrial DNA to psychiatric/behavioral phenotypes," in *BEHAVIOR GENETICS*, 2015, vol. 45, pp. 645–646.

[16] P. Geng, X. Tong, and Q. Lu, "An integrative U method for joint analysis of multi-level omic data," *BMC genetics*, vol. 20, no. 1, pp. 1–12, 2019.

[17] X. Zhang *et al.*, "Considering genetic heterogeneity in the association analysis finds genes associated with nicotine dependence," *Frontiers in genetics*, vol. 10, p. 448, 2019.

[18] Z. Zhu *et al.*, "Development of GMDR-GPU for gene-gene interaction analysis and its application to WTCCC GWAS data for type 2 diabetes," *PloS one*, vol. 8, no. 4, p. e61943, 2013.

[19] F.-T. Zhang, Z.-H. Zhu, X.-R. Tong, Z.-X. Zhu, T. Qi, and J. Zhu, "Mixed linear model approaches of association mapping for complex traits based on omics variants," *Scientific reports*, vol. 5, p. 10298, 2015.