**Yujue Wang, M.S.**

1853 Tomlinson Ave., Fl. 2, Bronx, NY, 10461; (347)-789-0342; [spsc83@gmail.com](mailto:spsc83@gmail.com)

**EDUCATION**

**East China University of Science and Technology**, *Shanghai, China*, **M.S.**, Control Science and Engineering, Sep. 2007 - May 2012

**Tongji University**, *Shanghai, China,* **B.S.**, Electrical Engineering and Automation, Sep. 2003 – Jun. 2007

**WORK EXPERIENCE**

**Albert Einstein College of Medicine, Genetics Department,** *NY, USA.*Apr. 2019-present, **Research Technician.**

As a research technician, I implement multiple projects and help the postdocs, PI, and labs with their projects.

**Project 1:** SomaMutDB is a website database that compiles all available somatic mutations data in healthy tissues including in-house data and publicly available data up to date. It provides multiple functions such as data visualization, data browsing, signature analyzing and data download. I implemented 6 tools including MutationalPatterns, SomaticSignatures, hdp, signature\_tools\_lib, Sigprofiler, and mmsig in the signature analyzing module of the website. I designed the structure of the website system, programmed the front end and the back end with Python, HTML, CSS, JavaScript, jQuery and etc. And I set up the MySQL database for the website. The paper of this website database was published on Nucleic Acids Research. I am the co-first author.

**Project 2:** PEA (phasing, enhanced reference genome and assembly) method is to identify genome structural variations from single-cell whole-genome sequencing (WGS) data. I designed the whole pipeline and programmed the pipeline with Python and Bash script under the supervision of a postdoc. The paper of PEA method was submitted to *Nature Methods*. I am the co-first author.

**Project 3:** I refactored and optimized a single-cell variant calling pipeline for one lab with Python. I simplified the whole pipeline from 200 commands and almost 800 parameters per cell to just 1 command and 4 parameters per cell, reduced the IO by 92.8%. And I sped up the pipeline by 10 times with parallel running technic. I also fixed the bugs in the previous version, and make it available for calling short insertions and deletions (INDELs).

**Project 4:** I expanded the DNA mapping pipeline’s availability from only for the human genome to 7 different species for one lab. And I implemented the pipeline for 48 samples of 6 different species for the lab.

**Project 5:** I annotated and classify SNVs and indels identified from the WGS data in >1500 22q11.2 deletion syndrome patients into damaging LoF, damaging frameshift, damaging Missense, benign Missense, splice-disrupting, synonymous variants using multiple state-of-art algorithms, software, and public databases including VEP, Bystro, spliceAI, etc. under the supervision of a postdoc. I built an SQLite database for it. And I implemented multiple statistical tests including, for example, Fisher’s exact test, binomial test, variant-Set Test for Association using Annotation infoRmation (STAAR) based on the database. The paper of the project is on the way and I am the second author.

**Project 6:** I built the lab websites for two PIs on the wowchemy and WIX platforms.

**Project 7:** I translated the global coordination level (GCL) analysis method from MATLAB to R for one postdoc.

**Project 8:** I developed a job submission website using Python, HTML, CSS, JavaScript and etc. for students and postdocs who are not familiar with Linux commands and high-performance computing clusters (HPC).

**Project 9:** I am developing a structural variation calling method for one PI with Python (ongoing).

**Albert Einstein College of Medicine**, **Genetics Department,** *NY, USA.*Jul. 2018-Apr. 2019, **volunteer and waiting for my H1b visa approval.**

As a volunteer I helped postdocs to implement their analyzing pipelines on HPC. And I learned necessary knowledge (single-cell WGS sequencing, DNA mapping, variant calling, etc.) for optimizing the SNV caller for single-cell sequencing data.

**China UnionPay Merchant Services Company, Ltd.**, Ufood Division, *Shanghai, China*. Dec. 2016-Nov. 2017, **Project Manager**.

Ufood provides solutions for restaurant management. As a project manager, I supervised 10 employees and managed the accounts of over 1000 different vendors. We designed applications for PC, tablet, and POS terminal, website for restaurant management, in addition to creating customized WeChat accounts for vendors and customers to facilitate reservations, ordering, and payments.

**China UnionPay Merchant Services Company, Ltd.**, MIS-POS (merchant integrated system – point of sale) Department, *Shanghai, China*. May 2012-Nov. 2016, **Programmer**.

I worked with senior personnel of international companies (Vanguard, McDonald's, Yum, NBA Play zone, Bestseller, ZARA, Decathlon, etc) to develop and implement their integrated payment system for the Chinese market. Our module is deployed on the computers of cashiers or vending machines to drive hardware and communicate with bank servers securely.

**SKILLS**

reading riding

Fluent in many coding languages: Python, R, Flask, MySQL, HTML, CSS, JavaScript, jQuery, Bash, C/C++, Delphi, Java, Perl, programming in Android and iOS environments

Mandarin Chinese (native speaker)

Proficient in Microsoft office suite; extensive PowerPoint presentation skills

Leadership experience

**PUBLICATIONS and WEBSITE DEVELOPMENTS**

Sun S\*, **Wang Y\***, Maslov AY, Dong X, Vijg J. SomaMutDB: a database of somatic mutations in normal human tissues. *Nucleic Acids Res.* 2022 Jan 7;50(D1): D1100-D1108. doi: 10.1093/nar/gkab914.

Dong X\*, Zhang L\*, Wang Y\*, Lee M., Maslov AY, Wang T, Gorbunova V, Vijg J. Identifying genome structural variations in single cells. In submission.

https://vijglab.einsteinmed.org/SomaMutDB/

<https://github.com/biosinodx/SCcaller>

<https://github.com/biosinodx/SCcaller3_PEA>

<https://www.vijglab.org/>

<https://srirajlab.com/>

**AWARDS AND HONORS**

Employee of the Year in 2014 and 2015, China UnionPay Merchant Services Company, Ltd.