Cheng-En Tan

The Tagkopoulos Lab, GBSF Rm 5312, 451Health Sciences Drive, UC Davis, CA, USA
(O) +1-530-752-7707
cetan@ucdavis.edu

OBJECTIVE: Bioinformatics scientist specialized in genomics and metabolomics data analysis in hopes of mining meaningful biological information and making contributions in systems biology.

EDUCATION:

University of California, Davis — Davis, California, USA
College of Engineering — In Progress

- Ph.D., Dept. of C.S. (GPA: 4.0/4.0)

National Taiwan University — Taipei, Taiwan
College of Electrical Engineering & Computer Science — November 2014

- Master of Science, Dept. of CSIE (GPA: 3.85/4.3)
- DISSERATION: *IDMass 2.0: Component Extraction, Identification and Statistics GUI Processing Toolkit for GC-MS* [2012-2014]

National Taiwan University — Taipei, Taiwan
College of Electrical Engineering & Computer Science — June 2012
College of Life Science

- **Double Degree of B.S.** (GPA: 87.77/100, The Dean Award (Top 10%))
  - Dept. of Biochemical Science and Technology (BST)
  - Dept. of Computer Science and Information Engineering (CSIE)

EXPERIENCE

University of California, Davis — Davis, California, USA
Research Assistant in The Tagkopoulos Lab — March 2022 – Present

- **Food sensory data curation and regression analysis**
- Predicting food flavor based on curated data using different models

3R Life Sciences Ltd. — Kaohsiung, Taiwan
Advanced Engineer — March 2020 – March 2022

- **Software development of Class III Medical Device**
  - Software and Firmware development of Ventricular Assist Device (VAD)

University of California, Davis — Davis, California, USA
Research Assistant in The Tagkopoulos Lab — September 2016 – March 2020

- **Microarray and RNA-seq data collection and analysis**
- Research of transfer learning on cross-species transcriptomic data

Process Integration and Predictive Analytics (PIPA LLC) — Davis, California, USA
Researcher Internship — July 2018 – August 2018

- **Implementation of RNA-seq data analysis systems using Python**
- Parallelize RNA-seq data analyzing process
- Performance evaluation of RNA-seq data analysis system
- Utilize RNA-seq data analysis tools including Bowtie2 and Htseq
- Parameters optimization of RNA-seq data analysis

National Taiwan University — Taipei, Taiwan
Research Assistant in Metabolomics Core Lab — November 2014 – July 2016

- **LC/GC-MS data analysis pipeline development using R**
- Graphical user interface implementation using wxPython package
- Integration of R code and Python code using RPy2 package
  - Conference Presentation: IDMass: GC/MS Data Management and Analyzes Software for Metabolomics Studies, Metabolomics Society, Japan, 2014

PUBLICATIONS

Peer-reviewed Article

1. Chan, T., **Tan, C. E.**, & Tagkopoulos, I. (2022). Audit lead selection and yield prediction from historical tax data using artificial neural networks. Plos one, 17(11), e0278121.
2. Kuo, T. C., **Tan, C. E.**, Wang, S. Y., Lin, O. A., Su, B. H., Hsu, M. T., ... & Tseng, Y. J. (2020). Human breathomics database. Database, 2020.
3. Tian, T. F., Wang, S. Y., Kuo, T. C., **Tan, C. E.**, Chen, G. Y., Kuo, C. H., ... & Tseng, Y. J. (2016). Web Server for Peak Detection, Baseline Correction, and Alignment in Two-Dimensional Gas Chromatography Mass Spectrometry-Based Metabolomics Data. Analytical chemistry, 88(21), 10395-10403.

Conference Presentation

1. **Tan, C. E.,** Chung, Y. Y., Kuo, C. H., Tseng, Y. J. (2014). IDMass: GC/MS Data Management and Analyzes Software for Metabolomics Studies, Metabolomics Society, Japan, 2014

ACADEMIC QUALIFICATION

- Master of Science, Department of Computer Science and Information Engineering, National Taiwan University. DISSERATION: *IDMass 2.0: Component Extraction, Identification and Statistics GUI Processing Toolkit for GC-MS* [2012-2014]
- **Double Degree of Bachelor of Science, Department of Biochemical Science and Technology and Department of Computer Science and Information Engineering**, National Taiwan University, with top 10% Grade Point Average and Dean Award, College of Life Science [2007-2012]

TOOLS

- **Programming language**: C, MATLAB, R and Python
- **Bioinformatics Software Design and Development:** Visual Studio
- **Application Programming Interface (API) Used:** Rpy2, WxPython
- **Database and API for access:** MySQL and Django

STRENGTHS

- Concept realization and implementation
- Desirable bridge between biologist and engineer

LINGUISTIC ABILITY

- Chinese: native language
- Cantonese: native language
- English: fluent (reading), intermediate (speaking, listening, writing)