Information

E-mail: yong.kweon@mail.mcgill.ca
Mailing address: Montreal, Quebec
Phone number: +1) 437-240-1343

Expertise

Multi-omics, Single cell, Bio-statistics/informatics, Statistical computing, Modeling and regression, Big data analysis, Machine learning

Education

MSc Thesis in Experimental Medicine (Jan 2023 ~)
McGill University, Montreal, Quebec, Canada
Supervisor: Jun Ding, PhD
Academic advisor: Janusz Rak, MD, PhD

PhD Candidate in Quantitative Life Science Withdrawn (Sep 2018 ~ Dec 2021)
McGill University, Montreal, Quebec, Canada
*Military service 2019 ~ 2021

BA in Statistics (Cluster: Mathematics/Economics) in May 2018
University of California, Berkeley, California, U.S.

AS-T in Mathematics and Business Administration & AA-T in Economics in May 2016
Diablo Valley College, Pleasant Hill, California, U.S.

Research

Cell type deconvolution Research under Professor Jun Ding **Jan 2023 – Present**
McGill University
● Decompose spatial transcriptomics (ST) at cell-type resolutions to gain comprehensive insights into spatial organization and uncover hidden biological information, such as cell functions and intricate interactions.
● Developed a method that uses deep neural network training on scRNA-seq and ST data, incorporating transfer learning from scRNA-seq reference data to ST data, and conducting iterative matrix factorizations.
● Implement a software that demonstrated higher accuracy, resolution, robustness, efficacy, and scalability on two simulated datasets and three real datasets, outperforming eight other popular and commonly used deconvolution methods.

AI Drug Design Research under Professor Jun Ding **Sep 2021 – Nov 2023**
McGill University
● Introduce scBeacon, an innovative framework built upon a VQ-VAE framework, deep contrastive siamese network, and a greedy iterative strategy, to effectively pinpoints differential genes and identify the same cell population across different biological conditions forming cluster pairs.
● Identify the common modified VQ-VAE structure for both Control and cytarabine drug (AraC) that consistently represents already-defined (Leiden UMAP initialized by PAGA) full-space clusters effectively.
● Annotate cluster cell types based on marker genes identified through the Wilcoxon test.
● Develop mapping networks between control clusters and AraC clusters using: 1) VAE with Zero-Inflated Negative Binomial distribution approximation, and 2) Contrastive learning.

● Analyze the expanded Connectivity Map (CMap) LINCS 2020 transcriptional datasets (level 5) using the cmapPy module to perform acute myeloid leukemia (AML) drug design analysis.

Statistical Bioinformatics Research under Professor Jeff Xia **Jan 2019 – May 2019**
McGill University
● Create a user-interactive website to display analysis outputs for multi-omics integration, including Two-way Orthogonal Partial Least Squares (O2PLS), Procrustes analysis, Multi-block Partial Least Squares, and Multi-block Principal Component Analysis.
● Develop and functionalize useful outputs of O2PLS, providing interpretations across various types of omics data such as proteomics, metabolomics, genomics, and transcriptomics.
● Benchmark DIABLO, mixOmics, and OmicsPLS, and extend the capabilities for multi-omics integration.
● Evaluate the quality and performance of O2PLS and n-way Orthogonal Partial Least Squares (OnPLS).

Neuroimaging Research under Professor Sylvain Baillet **Sep 2018 – Dec 2018**
Montreal Neurological Institute and Hospital, McGill University
● Develop an automated workflow pipeline for analysis on Brainstorm software (running on Matlab), including steps for artifact and noise removal, filter application, and statistical tools like entropy and power spectrum analysis.
● Perform data analysis for multimodal electrophysiology and imaging, such as MEG and EEG combined with MRI.
● Analyze data from a longitudinal study examining the "short-term" effects of brain stimulation with steady-state visual stimulation on auditory working memory performance in healthy volunteers.
● Conduct a new experiment investigating the "long-term" effects of brain stimulation with Transcranial Magnetic Stimulation (TMS) on auditory working memory performance in healthy volunteers, focusing on measurements with Electroencephalography (EEG).

Conjoint Analysis Research under Professor Philip Stark **Feb 2018 – May 2018**
University of California, Berkeley
● Demonstrate why conjoint analysis is logically flawed in market applications.

Research Intern under Professor Heather Haveman **Feb 2017 – Dec 2017**
University of California, Berkeley
● Collaborate on the project "Charter Schools and the Business Age" by manually and programmatically coding websites for text analysis to determine which charter schools thrive in the current political climate—those emphasizing discipline and college readiness or those prioritizing independent thinking and socio-emotional development.
● Work with Python/R coding teams to web-scrape and analyze the websites of all currently operating U.S. charter schools using seasonal time series analysis and machine learning techniques such as PCA and decision trees.
● Clean large datasets (dat, csv, and txt files) to make them publicly accessible and readable.

Research Intern under Professor Daniel Kammen **Sep 2016 – Dec 2016**
Renewable and Appropriate Energy Laboratory at University of California, Berkeley
● Contribute to the National Renewable Energy Deployment Research on the SWITCH project focusing on the Mexico region.
● Perform web scraping, data munging, and data visualization using PostgreSQL, Python, Pandas, and QGIS on hydropower data from 2006 to 2015, gaining insights into national energy policy.
● Clean large datasets from various sources and format them appropriately using non-parametric testing.
● Conduct various statistical analyses (hypothesis testing) and develop analytical predictions (time series) for modeling future hydro-plant power generation in Mexico.

● Wrote a scientific research paper on methods for forecasting future hydro-power generation.

Publications and Acknowledgements

Chenyu Liu, **Yong Jin Kweon**, Jun Ding (2023). scBeacon: single-cell biomarker extraction via identifying paired cell clusters across biological conditions with contrastive siamese networks. https://arxiv.org/abs/2311.02594

Haber, J.R. (2020). Sorting Schools: A Computational Analysis of Charter School Identities and Stratification. Sociology of Education. https://doi.org/10.1177/0038040720953218

Presentations and Posters

Albouy, P., Martinez-Moreno Z., **Kweon, Y.J.**, Zatorre, R.J., Baillet, S. (2018). Driving working memory with visual rhythmic stimulations. IDRC workshop. Montreal, Canada, Oct. 22, 2018.

Albouy, P., **Kweon, Y.J.**, Whittaker, H., Baillet, S., Zatorre, R.J. (2018). Enhancing learning-related plasticity with information-based neuromodulation. Auditory learning and plasticity symposium. Montreal, Canada, Oct. 23, 2018.

Teaching and Advising

Stat 133 TA/UGSI (Gaston Sanchez) **Jan 2018 – May 2018**
Statistics Department at University of California, Berkeley
● Lead 2-hour lab sections twice a week. (2 labs; 4 hours total)
● Hold 4-hour of office hours weekly.
● Proctor, make, grade midterm and final exams.
● Teach topics including statistical visualizations, analysis, web browser user interface with Shiny apps, regular expressions, reproducible R package development, Linux, Unix, Git, and more.

Stat 133 Lab Assistant (Adam Lucas, Gaston Sanchez) **Jan – May, Aug – Dec 2017**
Statistics Department at University of California, Berkeley
● Assist students with homework and lab assignments during lab hours.
● Provide tutoring in the R programming language.
● Gain experience in educating students on fundamental statistics topics.
● Enhance skills in critical thinking, communication, and teamwork.
● Grade R lab assignments according to the professor's rubric.

Haas Young Entrepreneurs at Haas (YEAH)/Boost Mentor **Sep 2015 – Dec 2016**
Haas School of Business at University of California, Berkeley
● Mentor "Group 1" consisting of five grade 9 students alongside two other Berkeley students. Provide guidance and share personal experiences to help them achieve their goals.
● Lead a group of grade 11 students in creating presentations advocating for the educational needs of students.

Professional Service

Apple Map Quality Data Analyst **Jan 2022 – Dec 2022**
Singapore Adecco Personnel Pte Ltd

Machine Learning Internship **May 2017 – Aug 2017**

IPMD

CFO/Co-Founder **Nov 2012 – June 2017**
Onbi, Smartphone Application Co

Marketing Internship **Oct 2014 – Jan 2015**
Transamerica

Skills and Languages

- Proficient in R.
- Intermediate in Python, SQL, and Latex.
- Familiar with MATLAB, Java, CSS, HTML, C++, Linux, Git, and Mathematica.
- Fluent in Korean and English. Studied Japanese, Chinese, and French (Creole).
- Proficiency in Windows and Microsoft Office (Word, Excel, and PowerPoint).
- Familiar with Photoshop.

Social Media

- **Website**

https://yjkweon24.wixsite.com/yongjinkweon
https://yjkweon24.github.io/

- **GitHub Projects**

https://github.com/yjkweon24/Jin-Projects-Show-up

- **LinkedIn**

https://www.linkedin.com/in/jin-kweon-5b687a103