YIMING LI

L1-05, Hong Kong Jockey Club Building for Interdisciplinary Research
5 Sassoon Road, Pokfulam, Hong Kong, Hong Kong
Mobile: (852) 65852837; **Email**: kestrel614@gmail.com; **Webpage**: https://kestrel614.github.io

EDUCATION AND QUALIFICATIONS

2014 – 2019 **University of Hong Kong**, Hong Kong
Ph.D. (Bioinformatics and Statistical Genetics)
Thesis title: Connecting the dots: Integrative analysis of genomic, metabolomic, and phenotypic data from a population cohort
Supervisors: Prof. Pak C. Sham (primary), Dr. Miaoxin Li and Dr. Stacey Cherny

2011 – 2014 **University of Hong Kong**, Hong Kong
B.Sc. (First Class Honours)
Major in Statistics / Minors in Computer Science & Mathematics
Cumulative GPA: 3.62; *Computer Science GPA*: 4.0

2013 Spring **University of California, Berkeley**, CA, U.S.A.
Reciprocal Exchange Student (Major in Statistics; GPA 4.0)

2010 – 2011 **Tsinghua University**, Beijing, P.R.China
Exchange Student (Major in Mathematical Science)

PUBLICATIONS

Desmond Campbell, **Yiming Li**, and Pak C. Sham. "Multifactorial disease risk calculator: Risk prediction for multifactorial disease pedigrees." *Genetic epidemiology* 42.2 (2018): 130-133.

Jian-Yu Shi, Siu-Ming Yiu, **Yiming Li**, Henry CM Leung, and Francis YL Chin. "Predicting drug-target interaction for new drugs using enhanced similarity measures and super-target clustering." *Methods* 83 (2015): 98-104. (Also presented in BIBM 2014 by **Yiming Li**)

PREPRINT

Yiming Li, Dino Samartzis, Jaro Karppinen, Kathryn S.E. Cheah, Danny Chan and Pak C. Sham. "Genomic and metabolomic analysis identify very low density lipoprotein as a potential risk factor for lumbar modic changes."

RESEARCH EXPERIENCE

2016 – present **Centre for Genomic Sciences, University of Hong Kong**, Hong Kong
Project Title: The Underlying Metabolomic Continuum of Lumbar Disc Degeneration
Supervisor: Prof. Pak C. Sham

- Integrating genomic, pheotypic and serum 1H NMR spectroscopy data for a population cohort of over 1000 individuals
- Reducing the dimension of metabolomic data via dynamic hierarchical tree cutting
- Performing genome-wide association studies (GWAS), conducting polygenic scoring based on the GWAS summary statistics and annotating the GWAS hits
- Analysing the association between clinical phenotypes and metabolomic polygenic scores via regression analysis and self-organizing maps
- Testing for causality of the detected associations using Mendelian randomization

RESEARCH EXPERIENCE (CONT'D)

2015 – present **Centre for Genomic Sciences, University of Hong Kong**, Hong Kong
Project Title: Longitudinal Study of MRI Features of the Human Lumbar Discs
Supervisor: Prof. Pak C. Sham
- Performing contingency table analysis (based on the log-linear model) on MRI feature data of the human lumbar intervertebral discs
- Fitting continuous-time structural equation models for insight into the etiology of lumbar disc degeneration
- Exploring and visualizing the underlying relationship structure of lumbar disc phenotypes through correlation networks and Markov network analysis
- Building a predictive model for future lumbar disc degeneration using random forests

2015 – 2016 **Centre for Genomic Sciences, University of Hong Kong**, Hong Kong
Project Title: Multifactorial Disease Risk Calculator: Web-Based Risk Prediction for Multifactorial Disease Pedigrees
Supervisors: Prof. Pak C. Sham; Dr. Desmond Campbell
- Devising an algorithm for estimating disease risk for pedigrees based on the liability-threshold model
- Implementing the algorithm in R (mainly) and C++ (Gibbs sampler)
- Creating a web interface for easy use of scientists and practitioners

2014 – 2015 **Centre for Genomic Sciences, University of Hong Kong**, Hong Kong
Project Title: Reviewing the Role of SNP Pre-Selection in the Polygenic Score Approach and Shrinkage Methods
Supervisors: Prof. Pak C. Sham; Dr. Timothy Mak
- Simulating dichotomous traits with different types of genetic architecture based on real genotype data
- Comparing the performances of different SNP pre-selection methods and risk prediction methods ranging from the polygenic score approach to LASSO
- Visualising the results in an interactive manner
- Results presented as a poster at the 2015 International Workshop on Statistical Genetic Methods for Human Complex Traits

2013 – 2014 **Department of Computer Science, University of Hong Kong**, Hong Kong
Project Title: Constructing the Drug-Protein Interaction Network
Supervisors: Prof. Francis Y.L. Chin; Dr. Jianyu Shi
- Crawling and integrating various types of drug-related data
- Devising a more accurate drug-protein interaction prediction model by employing machine learning methods

2013 – 2014 **Department of Statistics, University of Hong Kong**, Hong Kong
Project Title: Visualizing Big Ranking Data
Supervisor: Dr. Philip L.H. Yu
- Investigating possible improvements to the Spearman and Kendall distance formulas
- Developing a computationally efficient ranking data visualization framework
- Analyzing and visualizing several ranked datasets in R

2013 Spring **University of California, Berkeley**, CA, U.S.A.
Project Title: Supervised Learning: A Reliable Way to Predict U.S. Election Results?
- Predicting U.S. president election results using the k-nearest neighbors approach
- Visualizing the results using multi-dimensional scaling and hierarchical clustering
- Leading a team of five and providing constructive feedback to team members

RESEARCH EXPERIENCE (CONT'D)

2012 – 2013 **Centre for Genomic Sciences, University of Hong Kong**, Hong Kong
Project Title: Characterizing Genome-wide Complex Trait Analysis' Performance
Supervisors: Prof. Pak C. Sham; Dr. Desmond Campbell
- Evaluating the performance of genome-wide complex trait analysis software
- Simulating genome-wide association studies based on real genotype data
- Estimating the phenotypic variance explained by SNPs using the restricted maximum likelihood method

HONOURS AND AWARDS

2014 – 2018 HKU Postgraduate Fellowship

Nov 2014 The IEEE International Conference on Bioinformatics and Biomedicine (BIBM) Student Travel Award (25 awarded in total in BIBM 2014)

2011 – 2014 HKSAR Government Scholarship Fund
HKU Foundation Scholarships for Outstanding Mainland Students

2012 – 2013 HKU Worldwide Exchange Scholarship

2011 – 2012 Dean's Honours List (Top 10% of the class)

SKILLS

Programming Languages	Competent in R and Python; some knowledge of C++ and Java.
Operating Systems	UNIX / Linux, Mac OS X, Windows.
Professional Training	"Deep learning", a 5-course specialization on Coursera. [Certificate]
Languages	Mandarin Chinese (native), English (fluent).
Standardized Tests	*GRE (June 2011)*: V: 790 (99% Below), Q: 800 (94% Below) *TOEFL (Oct 2013)*: R: 29, L: 29, S: 26, W: 30

TEACHING AND NON-ACADEMIC WORK

Yiming Li and Yuqiong Li. "Deciphering the Trump tweets" (in Chinese). Popular science article related to text mining the tweets of Donald and Ivanka Trump, published via *Initium Lab* (2017). [Link]

Yiming Li. "Beginner's guide to R." Presenter of an introductory R workshop as part of *the Basic Research Skills Seminar Series*, Centre for Genomic Sciences, HKU (2017). [Materials]

Yiming Li and Philip Yu. "A picture is worth a thousand words" (in Chinese). Popular science article related to data visualisation, published in *The Hong Kong Economic Journal* (2014). [Link]

Yiming Li and Philip Yu. "Let the data speak" (in Chinese). Popular science article related to big data, published in *the Hong Kong Economic Journal* (2014). [Link]

ACTIVITIES AND INTERESTS

Conferences and Workshops: Poster presenter and participant of the 2015 International Workshop on Statistical Genetic Methods for Human Complex Traits. Presenter of the 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Participant of the 59^{th} World Statistics Congress and the 5^{th} International Congress of Chinese Mathematicians.

Volunteering: Instructor for English as a Second Language course to newly immigrated children in Kwai Hing, Hong Kong. Responsibilities include designing the curriculum, preparation and delivery of lectures, as well as coordinating teaching associated activities.

Personal: Interests include piano, Chinese martial arts and creative writing.