Li, Wei Vivian

683 Hoes Lane West Rm 210
Rutgers, The State University of New Jersey
Piscataway, NJ 08854

Telephone: 732-235-8574
Email: vivian.li@rutgers.edu
Homepage: http://liw-vivian.appspot.com/

RESEARCH INTERESTS

I am interested in statistical modeling and algorithm development, as well as their applications to high-throughput sequencing data in bioinformatics and clinical research. Major research topics of my lab include:

- Statistical modeling and experimental design of
  - Single-cell genomic data
  - Bulk-tissue high-throughput sequencing data
  - Biomedical data
- Gene regulatory networks
- Alternative RNA splicing
- Comparative bioinformatics

EDUCATION

University of California, Los Angeles (UCLA)   09/2014 - 06/2019
Department of Statistics
Ph.D. in Statistics
Dissertation: Statistical methods for bulk and single-cell RNA sequencing data
Advisor: Dr. Jingyi Jessica Li

Huazhong University of Science & Technology (HUST)   09/2010 - 06/2014
School of Mathematics and Statistics
B.S. in Statistics

EMPLOYMENT

Assistant Professor   08/2019 -
Department of Biostatistics and Epidemiology
School of Public Health
Rutgers, The State University of New Jersey

Teaching Assistant Consultant   10/2018 - 06/2019
Department of Statistics
University of California, Los Angeles

Graduate Student Researcher   09/2016 - 06/2019
Department of Statistics
University of California, Los Angeles

HONORS AND AWARDS

Busch Biomedical Grant Award, Rutgers	2020
The Caliburn Allocation Award of Research Computing, Rutgers	2020
The Most Outstanding Statistician Award, UCLA	2019
Physical Sciences Entrepreneurship & Innovation Award (co-inventor with Dr. Jessica Li), UCLA	2018
Dissertation Year Fellowship, UCLA	2018
Pearl Cohen Poster Award, UCLA Biomedical & Life Science Innovation Day	2018
Diversity Scholarship, UseR 2016 Conference	2016
Doctoral Student Travel Grant, UCLA	2016
The Most Promising Computational Statistician Award, UCLA	2015
Distinguished Graduate of Class 2014, HUST	2014
National Fellowship, Ministry of Education, China	2013
Student Award of Excellence (awarded to 20 students per year), HUST	2013
National Fellowship, Ministry of Education, China	2012
National Fellowship, Ministry of Education, China	2011
Outstanding Student Award (awarded to 3 students in each department), HUST	2011

RESEARCH FUNDING

Rutgers Busch Biomedical Grant   10/01/2020 - 09/30/2022
A Novel Approach for Joint Modeling of Multi-patient Single-cell Gene Expression Data

NJ ACTS Mini-Methods Grant   06/04/2020 - 06/03/2021
Statistical Methods for Single-cell Differential Gene Network Analysis

Rutgers School of Public Health Pilot Grant   11/01/2019 - 06/01/2021
Understanding Gene Co-expression Patterns in Single Cells

PUBLICATIONS

* co-first authors   __ corresponding authors   [†] trainees

Preprint:

1. **Li, W. V.**, Tong, X., Li, J. J. (2020). Bridging Cost-sensitive and Neyman-Pearson Paradigms for Asymmetric Binary Classification. *arXiv*.

2. Sun, T., Song, D., **Li, W. V.**, & Li, J. J. (2020). scDesign2: an interpretable simulator that generates high-fidelity single-cell gene expression count data with gene correlations captured. *bioRxiv*.

3. Roque, W., Cuevas-Mora, K., Sales, D., **Li, W. V.**, Rosas, I. O., & Romero, F. (2020). Transcriptomics, metabolomics and lipidomics of chronically injured alveolar epithelial cells reveals similar features of IPF lung epithelium. *bioRxiv*.

4. Jiang, R., **Li, W. V.**, & Li, J. J. (2020). mbImpute: an accurate and robust imputation method for microbiome data. *bioRxiv*.

Refereed Article:

5. <u>**Li, W. V.**</u> & Li, Y.† (2020). scLink: Inferring Sparse Gene Co-expression Networks from Single-cell Expression Data. *Genomics, Proteomics & Bioinformatics*, Accepted.

6. Zheng, Y., Chen Z., Han, L., Han, Y., Zou, X., Zhou, B., Hu, R., Hao J., Bai, S., Xiao H., **Li, W. V.**, Bueker, A., Ma Y., Xie, G., Yang, J., Chen S., Li, H., Cao, J., & Shen L. (2020). Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nature Communications*, 11(1), 1-17.

7. Wang, J., Deng, F., Zeng, F., Shanahan, A. J., **Li, W. V.**, & Zhang, L. (2020). Predicting long-term multicategory cause of death in patients with prostate cancer: random forest versus multinomial model. *American Journal of Cancer Research*, 10(5), 1344.

8. Xu, J., Cheng, Y., Yuan, X., **Li, W. V.**, & Zhang, L. (2020). Trends and prediction in daily incidence of novel coronavirus infection in China, Hubei Province and Wuhan City: an application of Farr's law. *American Journal of Translational Research*, 12(4), 1355.

9. **Li, W. V.**∗, Li, S.∗, Tong, X., Deng, L., Shi, H., & Li, J. J. (2019) AIDE: annotation-assisted isoform discovery with high precision. *Genome Research*, 29, 2056-2072.

10. **Li, W. V.** & Li, J. J. (2019). A statistical simulator scDesign for rational scRNA-seq experimental design. *Bioinformatics (ISMB/ECCB 2019)*, 35(14), i41-i50.

11. Ge, X., Zhang, H.†, Xie, L., **Li, W. V.**, Kwon, S. B., & Li, J. J. (2019). EpiAlign: an alignment-based bioinformatic tool for comparing chromatin state sequences. *Nucleic Acids Research*, gkz287.

12. **Li, W. V.** & Li, J. J. (2018). An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nature Communications*, 9, 997.

13. **Li, W. V.** & Li, J. J. (2018). Modeling and analysis of RNA-seq data: a review from a statistical perspective. *Quantitative Biology*, 6(3), 195-209.

14. **Li, W. V.**, Zhao, A., Zhang, S., & Li, J. J.† (2018). MSIQ: joint modeling of multiple RNA-seq samples for accurate isoform quantification. *Annals of Applied Statistics*, 12(1), 510-539.

15. **Li, W. V.**, Chen, Y., & Li, J. J. (2017). TROM: A testing-based method for finding transcriptomic similarity of biological samples. *Statistics in Biosciences*, 9(1), 105-136.

16. **Li, W. V.**, Razaee, Z. S., & Li, J. J. (2016). Epigenome overlap measure (EPOM) for comparing tissue/cell types based on chromatin states. *BMC Genomics*, 17(1), S10.

Dissertation:

17. **Li, W. V.** (2019). Statistical methods for bulk and single-cell RNA sequencing data. PhD dissertation, UCLA.

Google Scholar: <https://scholar.google.com/citations?user=HzjANkQAAAAJ>

INVITED TALKS AND CONFERENCE PRESENTATIONS

Conference and Symposium Presentations:

1. Invited Talk, Sixth International Conference on Statistics for Twenty-first Century 12/2020
   Statistical Challenges in Single-cell RNA Sequencing Data Analysis

2. Poster, RECOMB/ISCB 2020 Conference on Regulatory and Systems Genomics   11/2020
   Inferring Gene Co-expression Networks from Single Cell Gene Expression Data

3. Talk, Human Cell Atlas Asia Meeting 2020   10/2020
   Inferring Gene Co-expression Networks from Single Cell Gene Expression Data

4. Talk, 2020 Joint Statistical Meetings, Virtual.   08/2020
   AIDE: Annotation-Assisted RNA Transcript Discovery with High Precision

5. Invited Talk, RCAF Forum: Challenges and Opportunities in   08/2020
   Computational Biology.
   Current Computational and Analytical Challenges in Transcriptomics

6. Poster, 28th Conference on Intelligent Systems for Molecular Biology   07/2020
   (ISMB 2020).
   Inferring Gene Co-expression Networks from Single Cell Gene Expression Data

7. Invited Talk, RCAF Webinar Series.   05/2020
   Statistical Challenges and Opportunities in Single-cell Studies

8. Invited Talk, ENAR 2020 Spring Meeting, Virtual.   03/2020
   A Unified View of Asymmetric Binary Classification

9. Invited Talk, Inaugural Lange Symposium Software Workshop   02/2020
   Improving Single-cell RNA-seq Data Analysis with scImpute and scDesign

10. Invited Talk, 2019 Joint Statistical Meetings, Denver, CO.   07/2019
    A Unified View of Asymmetric Binary Classification

11. Invited Talk, ENAR 2019 Spring Meeting, Philadelphia, PA.   03/2019
    A Unified View of Asymmetric Binary Classification

12. Talk, 8th Annual Southern California Systems Biology Conference, Irvine, CA.   02/2019
    AIDE: Annotation-assisted Isoform Discovery with High Precision

13. Invited Talk, 2018 IEEE 8th International Conference on Computational   10/2018
    Advances in Bio and Medical Sciences, Las Vegas, NV.
    scImpute: Accurate and Robust Imputation for Single Cell RNA-seq Data

14. Talk and Poster, Inaugural Symposium on Multiscale Cell Fate, Irvine, CA.   10/2018
    scImpute for Imputation of ScRNA-seq Data

15. Talk, UCLA 4th Annual QCBio Retreat, Los Angeles, CA.   09/2018
    AIDE for Annotation-assisted mRNA Isoform Reconstruction and Quantification

16. Talk, 2018 Joint Statistical Meetings, Vancouver, Canada.   08/2018
    Joint Modeling of Multiple RNA-seq Samples for Accurate Isoform Quantification

17. Talk, 26th Conference on Intelligent Systems for Molecular Biology   06/2018
    (ISMB 2018), Chicago, IL.
    scImpute: Accurate and Robust Imputation for Single Cell RNA-seq Data

18. Poster, UCLA Biomedical & Life Science Innovation Day, Los Angeles, CA.   06/2018
    scImpute: Accurate Imputation for Single Cell RNA-seq Data

19. Invited Talk, DahShu Webinar.   04/2018
    An Accurate and Robust Imputation Method for Single-cell RNA-seq Data

20. Invited Talk, 3rd Annual NGS Data Analysis & Informatics Conference,   02/2018
    San Diego, CA
    scImpute: Accurate and Robust Imputation for Single Cell RNA-seq Data

21. Poster, UCLA 3rd Annual QCBio Retreat, Los Angeles, CA.   09/2017
    scImpute: Accurate and Robust Imputation for Single Cell RNA-seq Data

22. Talk, Asia Pacific Bioinformatics Conference 2016, San Francisco, CA.   01/2016
    Epigenome Overlap Measure (EPOM) for Comparing Tissue/cell Types based on
    Chromatin States

23. Talk, USC/UCLA Bioinformatics Joint Meeting, Los Angeles, CA.   11/2015
    A Testing-based Measure EPOM for Comparing Tissue/cell Types based on
    Chromatin States

24. Poster, UCLA 1st Annual QCBio Retreat, Los Angeles, CA.   09/2015
    A New Measure for Comparing Epigenetic Profiles in Tissue/cell Types

    Seminar Presentations:

25. Epigenetics User Group Meeting, Rutgers Department of Genetics   02/2020
    Improving Statistical Analysis of Bulk and Single-cell RNA-seq Data

26. Systems Biology/Bioinformatics Meeting, Rutgers Cancer Institute of New Jersey 10/2019
    Improving Statistical Analysis of Bulk and Single-cell RNA-seq Data

27. Department of Biostatistics and Epidemiology, Rutgers University.   02/2019
    Statistical Methods to Uncover Hidden Information from Large-scale Genomic Data

28. Department of Statistics, University of Wisconsin-Madison.   02/2019
    Statistical Methods to Uncover Hidden Information from Large-scale Genomic Data

29. Department of Public Health, University of California, Merced.   01/2019
    Statistical Methods to Uncover Hidden Information from Large-scale Genomic Data

30. Department of Statistics, Pennsylvania State University.   01/2019
    Statistical Methods to Uncover Hidden Information from Large-scale Genomic Data

31. Department of Statistics, University of Illinois at Urbana-Champaign.   12/2018
    Statistical Methods to Uncover Hidden Information from Large-scale Genomic Data

32. Jonsson Comprehensive Cancer Center Gene Regulation Intramural   03/2018
    Meeting, UCLA.
    scImpute: Accurate and Robust Imputation for Single Cell RNA-seq data

TEACHING EXPERIENCE

Instructor (Rutgers Graduate Courses):

BIST 0725: Generalized Linear Models Spring 2020

Student Instructor (UCLA Graduate Courses):

STATS 495A: Teaching College Statistics Fall 2018

Teaching Assistant (UCLA Graduate Courses):

STATS 207: Statistical Learning with Sparsity Spring 2018

STATS 203: Large Sample Theory, Including Resampling Winter 2017

STATS 205: Hierarchical Linear Models Winter 2017

STATS M254: Statistical Methods in Computational Biology Spring 2016

STATS 200C: Large Sample Theory Spring 2016

Teaching Assistant (UCLA Undergraduate Courses):

STATS 101A: Introduction to Data Analysis and Regression Summer 2016

STATS 13: Introduction to Statistical Methods for the Health and Life Sciences Winter 2016

STATS 10: Introduction to Statistical Reasoning Fall 2015

STATS 100B: Introduction to Mathematical Statistics Summer 2015

Guest Lecturer:

RESH 0701: Doctoral Colloquium 11/2020
Rutgers School of Public Health
Career As a Statistician in the Bioinformatics Field

DSO 530: Applied Modern Statistical Learning Methods 04/2020
Marshall School of Business, University of South California
Introduction to Cloud Computing

DSO 530: Applied Modern Statistical Learning Methods 03/2019
Marshall School of Business, University of South California
Cloud Computing with the Amazon Web Services

BIOINFO 201: Bioinformatics Interdisciplinary Research Seminar 10/2018
Bioinformatics Interdepartmental Ph.D. Program, UCLA
Statistical Methods for Isoform Quantification of Multiple Samples

STATS M254: Statistical Methods in Computational Biology 06/2018
Department of Statistics, UCLA
Statistical Methods for Isoform Quantification of Multiple Samples

STATS M254: Statistical Methods in Computational Biology 05/2018
Department of Statistics, UCLA
Imputation Methods for Single-cell RNA-seq Data

DSO 530: Applied Modern Statistical Learning Methods 02/2018
Marshall School of Business, University of South California
Cloud Computing with the Amazon Web Services

DSO 530: Applied Modern Statistical Learning Methods 03/2017
Marshall School of Business, University of South California
Introduction to the Amazon Web Services (AWS)

STUDENT MENTORING & ADVISING

Research Mentor of Graduate Students:

- Dayuan Wang (2020/08 -) MS student in Biostatistics at Rutgers
- Shiwei Fu (2020/06 -) MS student in Statistics at UIUC
- Alireza Naghizadeh (2020/06 - 2020/08) PhD student in Computer Science at Rutgers
- Yanzeng Li (2019/12 - 2020/09) MS student in Biostatistics at Rutgers

Research Mentor of Undergraduate Students:

- Andrew Cheng (09/2020 -) Undergraduate student in Computer Science at Rutgers

Academic Advisor of Graduate Students:

- Phoebe Chu (08/2020 -) MS student in Biostatistics at Rutgers
- Changhan Lin (08/2020 -) MS student in Biostatistics at Rutgers
- Adriel Staana (08/2020 -) MS student in Biostatistics at Rutgers
- Daniel Yen (08/2020 -) MS student in Biostatistics at Rutgers
- Zifan Guo (01/2020 -) MS student in Biostatistics at Rutgers

Thesis Committee Member:

- Wenxuan Xiong (2020 -) PhD student in Biostatistics at Rutgers
- Aysegul Guvenek (2019 - 2020) PhD student in Biomedical Science at Rutgers

PROFESSIONAL SERVICE

Committee Member:

Rutgers School of Public Health Bylaws and Elections Committee 2020 -
Rutgers Bioinformatics and Microbiome Research Committee 2020 -

Reviewer:

Applied Sciences ($\times 1$)
Bioinformatics ($\times 8$)
BioMed Research International ($\times 1$)
Biometrics ($\times 1$)
BMC Bioinformatics ($\times 1$)
BMC Genomics ($\times 1$)
Cells ($\times 1$)
Chapman & Hall/CRC Press ($\times 1$)
Frontiers in Genetics ($\times 2$)
Gene Report ($\times 1$)
Genes ($\times 2$)

Genome Biology (×1)
Genomics, Proteomics & Bioinformatics (×1)
IEEE Access (×2)
Journal of Molecular Cell Biology (×1)
Laboratory Investigation (×2)
NAR Cancer (×1)
NAR Genomics and Bioinformatics (×1)
Nature Communications (×1)
Nature Methods (×1)
Nucleic Acids Research (×3)
PeerJ (×1)
Plos Computational Biology (×3)
Plos One (×1)
Proteomics (×1)
Quantitative Biology (×1)

PROFESSIONAL AFFILIATIONS

Member:

Rutgers Cancer Institute of New Jersey	2019 -
Eastern North American Region International Biometric Society	2019 -
American Statistical Association	2019 -
International Chinese Statistical Association	2019 -

Student Member:

International Society for Computational Biology	2017 - 2019
International Chinese Statistical Association	2015 - 2019
American Statistical Association	2014 - 2019

SOFTWARE

R packages:

· **scDesign** for rational scRNA-seq experimental design
  https://github.com/Vivianstats/scDesign

· **scLink** for constructing single-cell gene co-expression networks
  https://github.com/Vivianstats/scLink

· **scImpute** for accurate imputation of scRNA-seq data
  https://github.com/Vivianstats/scImpute

· **AIDE** for annotation-assisted transcript reconstruction and quantification
  https://github.com/Vivianstats/AIDE

· **MSIQ** for quantifying expression of RNA transcripts from multiple RNA-seq samples
  https://github.com/Vivianstats/MSIQ

· **TROM** for comparing transcriptomes of biological samples from the same or different species
  https://cran.r-project.org/web/packages/TROM/index.html