

# Yonghua Zhuang, PhD, PhD

Assistant Professor

Biostatistics and Bioinformatics Shared Service at Cancer Center

University of Colorado Denver Anschutz Medical Campus

Education 2 South Room 5323

13121 E. 17th Ave.

Aurora, Colorado 80045

E-mail: [Yonghua.Zhuang@CUAnschutz.edu](mailto:Yonghua.Zhuang@CUAnschutz.edu)

## Experience

---

**University of Colorado AMC, Aurora CO** Assistant Professor **Mar 2022-Present**

- Providing quantitative support for the planning, design, analysis, and presentation of basic science, clinical, and epidemiological investigations.
- Building an automated pipeline for single cell RNAseq analysis.
- Developing statistical methods for multi-omics integration.
- Assisting grant applications and manuscript preparations.

**NHLBI BioData Catalyst Fellow** **Sep 2020-Jan 2022**

- Developed an augmented high-dimensional Graphical Lasso model to incorporate prior biological knowledge for global network learning.
- Identified miRNA-mRNA Networks associated with COPD Phenotypes.

**Arrow Electronics, Denver CO** Data Scientist Intern **Jun 2019-Aug 2019**

- Developed several machine learning models including Supplier Requesting Quote (SQR) outcome classification and Slow-Moving Inventory (SMI) prediction.
- Deployed one prediction model with TensorFlow Serving for upcoming business production.
- Participated in an intense case competition for protecting security and privacy. Won 1st place with teammates for developing a comprehensive solution incorporating Artificial Intelligence (AI) technology that secures Internet of Medical Things (Iomt).

**University of Colorado Denver AMC, Aurora CO** Research assistant **Sep 2017-Feb 2022**

- Developing deep graph neural network method to identify gene networks associated to COPD disease.
- Developed a novel method for unsupervised discovery of phenotype-specific multi-omics networks.
- Developed a tissue augmented Bayesian model for expression quantitative trait loci analysis.
- Engineered proteomic features with Autoencoder and developed dense neural network for COPD classification with 78% accuracy (COPDGene Deep Learning workshop, Denver, 2018).
- Developed [Bayesian Deep classifier](#) (No-U-Turn Hamiltonian Monte Carlo Sampler) for Parkinson's disease using voice data. The F1 score was 0.82.

**Insight Health Data Science, Boston MA** Fellow **May 2017-Sep 2017**

- Built an automated acute Myeloid Leukemia (AML) classifier ([aml-classifier](#)), a machine learning-powered system that classifies whether a person has AML with flow cytometry (laser-based technology to analyze cell characteristics) data. Solved the issues of labor-intensive manual analysis and low reproducibility for high dimension flow cytometry data.
- Balanced dataset with sub-sampling, split dataset into training and test datasets and implemented random forest and SVM to train and predict AML with 94% accuracy.

**University of Colorado Denver AMC, Aurora CO** Research instructor/Fellow **Jan 2010-Sep 2017**

- Developed augmented Bayesian methods to integrate mouse lung eQTL (associations between gene expression and SNPs, Single-nucleotide polymorphisms) (40G raw data) to improve the prediction of liver eQTL by 8%. Demonstrated that the newly developed method is more robust than existing methods especially when the dataset is limited.
- Identified 59 significant cellular genes and 9 signaling pathways that are dysregulated in the mouse brain following infection with Zika virus with Next generation RNA sequencing data (10G raw data).
- Profiled microarray gene expression in idiopathic pulmonary arterial hypertension (IPAH), performed PCA and clustering analyses. Performed Student's test with BH correction and found 245 significantly differentially expressed genes.
- Investigated causal effect of surgery in conjunction with radiation on the survival of lymphoma patients in potential outcome framework. Determined propensity score using non-parsimonious logistic regression and linear discriminant analysis (LDA). Determined causal effect with propensity score matching and performed sensitivity analysis.
- Utilized Python Twitter API to retrieve tweeter posts using “Zika virus” keyword, investigated the pattern of Zika virus tweets using NLP and identified geographic distribution of tweets.

## Education

|   |                  |
|---|------------------|
| <b>Ph.D in Biostatistics, University of Colorado AMC, Aurora, CO (GPA 4.0/4.0)</b>        | <b>Feb 2022</b>  |
| <b>M.S. in Biostatistics, University of Colorado Denver AMC, Aurora, CO (GPA 4.0/4.0)</b> | <b>Dec 2016</b>  |
| <b>Ph.D. in Immunology, Sichuan University, Chengdu, China</b>                            | <b>July 2009</b> |
| <b>BSc in Clinical Medicine (MD), Tongji University, Shanghai, China</b>                  | <b>July 2001</b> |

## Selected Publications

- [1] Y. Zhuang, F. Xing, D. Ghosh, F. Banaei-Kashani, R. Bowler, K. Kechris. An augmented high-dimensional Graphical Lasso method to incorporate prior biological knowledge for global network learning. *Frontiers in Genetics*. 2021 (Accepted).
- [2] Y. Zhuang, BD. Hobbs, C. Hersh, K. Kechris. Identifying miRNA-RNA Networks associated with COPD Phenotypes. *Frontiers in Genetics*. 2021 Oct 28;12:748356.
- [3] L. Gillenwater, S. Helmi, E. Stene, K. Pratte, Y. Zhuang, R. Schuyler, L. Lange, P. Castaldi, C. Hersh, F. Banaei-Kashani, R. Bowler, K. Kechris. Multi-Omics Subtyping Pipeline for Chronic Obstructive Pulmonary Disease. *Plos One*. 2021 (Accepted).
- [4] E. Mastej, L. Gillenwater, Y. Zhuang, KA. Pratte, RP. Bowler, K. Kechris. Identifying Protein-metabolite Networks Associated with COPD Phenotypes. *Metabolites*. 2020 Mar 25;10(4):124.
- [5] Y. Zhuang, K. Wade, LM. Saba, K. Kechris. Development of a tissue augmented Bayesian model for expression quantitative trait loci analysis. *Math Biosci Eng*. 2019 Sep 26;17(1):122-143.
- [6] WJ. Shi, Y. Zhuang, PH. Russell, BD. Hobbs, MM. Parker, PJ. Castaldi, P. Rudra, B. Vestal, CP. Hersh, LM. Saba, K. Kechris. Unsupervised discovery of phenotype-specific multi-omics networks. *Bioinformatics*. 2019 Nov 1;35(21):4336-4343 (co-first author).
- [7] P. Clarke P, Y. Zhuang, HM. Berens, JS. Leser, KL. Tyler. Interferon Beta Contributes to Astrocyte Activation in the Brain following Reovirus Infection. *J Virol*. 2019 May 1;93(10):e02027-18 (co-first author).
- [8] E. Halper-Stromberg, L. Gillenwater, C. Cruickshank-Quinn, WK. O'Neal, N. Reisdorph, I. Petrache, Y. Zhuang, WW. Labaki, JL. Curtis, J. Wells, S. Rennard, KA. Pratte, P. Woodruff, KA. Stringer, K. Kechris, RP. Bowler. Bronchoalveolar Lavage Fluid from COPD Patients Reveals More Compounds Associated with Disease than Matched Plasma. *Metabolites*. 2019 Jul 25;9(8):157.

- [9] Y. Zhuang, H. M. Berens-Norman, J. S. Leser, P. Clarke and K. L. Tyler. Mitochondrial p53 contributes to reovirus-induced neuronal apoptosis and CNS injury in a mouse model of viral encephalitis. *Journal of Virology* 2016 Jun 15. pii: JVI.00583-16.
- [10] X. Qi, J. Hong, Lee Chaves, Y. Zhuang, Y. Chen, D. Wang, J. Chabon, B. Graham, K. Ohmori, Y. Li, and H. Huang. Antagonistic regulation by C/EBP $\alpha$  and MITF specifies basophil and mast cell fates. *Immunity*. 2013 Jul 25;39(1):97-110.
- [11] Z. Chen, S. Wang, N. Erekosima, Y. Li, J. Hong, X. Qi, P. Merkel, V. Nagabhushanam, E. Choo, R. Katial, R. Alam, A. Trikha, H. Chu, Y. Zhuang, M. Jin, C. Bai, H. Huang. IL-4 confers resistance to IL-27-mediated suppression on CD4+ T cells by impairing signal transducer and activator of transcription 1 signaling. *J Allergy Clin Immunol*. 2013 Oct;132(4):912-21.
- [12] J. Nishida, Y. Li, Y. Zhuang, Z. Huang, H. Huang. IFN-gamma suppresses permissive chromatin remodeling in the regulatory region of the Il4 gene. *Cytokine*. 2013 Mar 9. pii: S1043-4666(13)00063.
- [13] K. Dionne, Y. Zhuang, J. Leser, K. Tyler, and P. Clarke. Daxx upregulation in the cytoplasm of reovirus-infected cells is mediated by interferon and contributes to apoptosis. *Journal of Virology*, 2013 Mar;87(6):3447-60. doi: 10.1128/JVI.02324-12.
- [14] Y. Zhuang, S. Durrani, B. Hodges, S. C. Dreskin and X. Chen. Expression of recombinant Ara h 6 in *Pichia pastoris* but not in *E. coli* preserves allergic effector function and allows assessment of specific mutations. *Molecular Nutrition & Food Research*, 2012, 56(6): 986-995.
- [15] Y. Zhuang, S. C. Dreskin. Redefining the major peanut allergens. *Immunologic Research*. 2013 Mar;55(1-3):125-34. doi: 10.1007/s12026-012-8355.
- [16] X. Chen, Q. Wang, R. El-Mezayen, Y. Zhuang, and S. C. Dreskin. Ara h 2 and Ara h 6 have similar allergic effector activity and are substantially redundant. *International Archives of Allergy and Immunology*, 2012 Oct 16;160(3):251-258.
- [17] N. Goplen, M. Z. Karim, L. Guo, Y. Zhuang, H. Huang, M. M. Gorska, E. Gelfand, G. Pagés, J. Pouysségur, and R. Alam. Role of ERK1 in a Mouse Model of Chronic and Acute Asthma. *FASEB*, 2012, 26(5):1934-45.
- [18] M. Kulis, X. Chen, J. Lew, Q. Wang, O. P. Patel, Y. Zhuang, H. S. Porterfield, A. Burks, and S. C. Dreskin. The 2S albumin allergens of *Arachis hypogaea*, Ara h 2 and Ara h 6, are the major 2 elicitors of anaphylaxis and can effectively desensitize peanut-allergic mice. *Clinical and Experimental Allergy*, 2012, 2012, 42(2):326-36.
- [19] X. Chen, Y. Zhuang, Q. Wang, D. Moutsoglou, G. Ruiz, S.-E. Yen and S. C. Dreskin. Analysis of the effector activity of Ara h 2 and Ara h 6 by selective depletion from a crude peanut extract. *Journal of Immunological Methods*. 2011, Sep 30;372(1-2):65-70.
- [20] Y. Zhuang, Z. Huang, J. Nishida, M. Brown, L. Zhang, and H. Huang. A continuous T-bet expression is required to silence the interleukin-4-producing potential in T helper type 1 cells. *Immunology*. 2009, 128:34-42.
- [21] Y. Zhuang, Z. Huang, J. Nishida, M. Brown, L. Zhang, and H. Huang. Signaling pathways that lead to the silencing of the interleukin-4-producing potential in Th1 cells. *J Interferon Cytokine Res*. 2009, 29:399-406.

## HONORS, GRANTS AND AWARDS

---

- 2011, AAAAI Fellows-in-Training travel scholarship
- 2012, AAAAI Fellows-in-Training travel scholarship
- 2015, ISNV Investigators in Training scholarship
- 2020, NHLBI BioData Catalyst Fellowship

- 2021, Strother Walker Award for Outstanding PhD Student

## Software development

---

- SmCCNet R package development and maintenance
- AhGlasso R package development (to be submitted)

## Skills

---

**Languages:** Python (5 years), R (8 years), SAS (5 years), SQL (2 years), bash scripting, Java (Basic), Javascript  
**Tools:** Scikit-learn, Pandas, Numpy, Scipy, Tensorflow, Keras, Pytorch, Flask, R Caret, R ggplot2, R Shiny, Git, Amazon Web Service (AWS), Google Cloud, Docker

**Machine Learning and Statistics:** Deep learning, Graph Neural Network, KNN, SVM, LSTM, Autoencoder, Random Forest, Clustering analysis, Sensitivity analysis, Survival analysis, Causal inference, Longitudinal data analysis, Generalized mixed model, Bayesian inference, PCA, Stochastic modeling, Natural Language Processing (NLP), Data simulation

**Consulting:** Statistics consulting for COPDgene clinical study (5 years)

## REVIEW AND REFEREE WORK

---

Ad hoc Reviewer for *Bioinformatics*, *BMC bioinformatics*, *Journal of Virology*, *Plos One*, *Infection and Immunity*, *Journal of Hospital Infection*, *Cytokine*, *International Immunopharmacology*, *Frontiers in Oncology*.