CURRICULUM VITAE

Ruifeng Hu, Ph.D.

Current institute: Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston (UTHealth-Houston), Houston, Texas, USA

Telephone: 1 (979) 264-7270

Email: huruifeng.cn@hotmail.com / ruifeng.hu@uth.tmc.edu

Research interests and skills

My research is mainly to develop methods to unveil the hidden biological circuitries underlying the data, from understanding sequence-based regulations to the genomes or proteome and their impacts to diseases, including: (i) Decoding genomic sequence to decipher connections between sequence, functions, and phenotypes, especially leveraging sequence data and deep learning models; (ii) Developing new approaches for prioritizing actionable mutations characterized human diseases; (iii) Investigating gene expression patterns in human traits/diseases utilizing (epi)genomics data; (iv) Data integration, database construction, webserver platform development; (v) Multi-omics data analysis and a variety of bioinformatics statistical modeling.

- ➢ Proficient in Python, C++, and Linux shell language and large-scale data calculation
- ➢ Proficient in various algorithms of deep learning, including CNN, RNN, DNN, GAN, Autoencoder and using them to develop prediction algorithms.
- ➢ Proficient in common clustering and classification algorithms of machine learning, including naive bayesian classifier, SVM, k-means, decision tree, random forest.
- ➢ Proficient in construction of bioinformatics webserver platforms and databases (HTML, Python, PHP, JS).
- ➢ Proficient in various bioinformatics statistical modeling and data analysis software.
- ➢ Rich experiences in various types of NGS data such as RNA-seq, ChIP-seq, ATAC-seq, et al.
- ➢ Experience in the analysis of multi-omics data and single cell data.
- ➢ Familiar with molecular biology techniques and experiments.

EXPERIENCE

2018.4 – Present   Postdoctoral Research Fellow, Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA

EDUCATION

2012.9 – 2017.7   Ph.D. Pharmacognosy, with a concentration in Bioinformatics

Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences & Peking Union Medical College (CAMS&PUMC), Tsinghua University, Beijing, China

2008.9 – 2012.6   B.S. Computer Science and Technology
College of Computer Science and Technology, Nanjing Forestry University (NJFU). Nanjing, China

JOURNAL PEER REVIEW

Bioinformatics, Database, BMC Supplements, Frontiers in Genetics, PLoS ONE, International Conference on Intelligent Biology and Medicine 2020, Scientific Reports.

DATABASES/TOOLS/WEBSERVERS

1. Dr.VAEN: Drug response prediction using Variational Autoencoder based Elastic Net models. https://bioinfo.uth.edu/drvaen .
2. MitoX: exploring mitochondrial heteroplasmy and gene expression from single-cell sequencing assays. https://github.com/huruifeng/MitoX
3. CSEA-DB: Cell-type-Specific Enrichment Analysis DataBase. https://bioinfo.uth.edu/CSEADB
4. KinaseMD: Kinase mutations and drug responses. https://bioinfo.uth.edu/kmd
5. DeepFun: A tissue and cell type specific deep learning sequence-based model to decipher noncoding variant effects. https://bioinfo.uth.edu/deepfun
6. 6mA-Finder: Computational prediction of 6mA sites in the DNA sequences. https://bioinfo.uth.edu/6mA_Finder
7. TSEA-DB: A trait-tissue association map for human complex traits and diseases. https://bioinfo.uth.edu/TSEADB
8. VISDB: A knowledgebase containing the most comprehensive DNA virus integration sites and also the integration sites for several typical RNA retrovirus. https://bioinfo.uth.edu/VISDB
9. CleftGeneDB: The cleft gene annotation database aiming to provide a resource or reference for the midfacial development and the related disease studies for mouse and human. https://bioinfo.uth.edu/CleftGeneDB
10. ANCO-GeneDB: annotations and comprehensive analysis of candidate genes for alcohol, nicotine, cocaine and opioid dependence. https://bioinfo.uth.edu/ancogenedb
11. TarNet: an evidence-based database for natural medicine research. https://github.com/huruifeng/TarNet
12. LSCplus: a fast solution for improving long read accuracy by short read alignment. https://github.com/huruifeng/LSCplus
13. lncRNATargets: A platform for lncRNA target prediction based on nucleic acid thermodynamics. https://github.com/huruifeng/lncRNATarget_local

PUBLICATIONS

Manuscript in review/revision

1. **Ruifeng Hu**, Zhongming Zhao*. MitoX: exploring mitochondrial heteroplasmy and gene expression from single-cell sequencing assays. *Bioinformatics*. (Under review).

2. Hao-Dong Xu, Fangfang Yan, **Ruifeng Hu**, Akiko Suzuki, Chihiro Iwaya, Peilin Jia, Junichi Iwata*, Zhongming Zhao*. CleftGeneDB. a manually curated cleft-related gene resource. *Briefings in Bioinformatics*.(Under review).
3. Peilin Jia[#], **Ruifeng Hu[#]**, Guangsheng Pei, Yulin Dai, Yinying Wang, Zhongming Zhao. Deep generative neural network based drug response imputation reveals accurate identification of genetic and genomic signatures. *Nature Communication* (<u>Co-first author,</u> manuscript in minor revision).
4. Jian Tu, Zijun Huo, Yao Yu, Dandan Zhu, **Ruifeng Hu**, An Xu, Ruoyu Wang, Haidan Luo, Julian A. Gingold, Jie Su, Mo-Fan Huang, Kuang-Lei Tsai, Ruoji Zhou, Hui-Ming Chen, Weiling He, Shu-Hsia Chen, Thomas R. Webb, Huiling Yang, Peilin Jia, Jason T. Yustein, Lisa L. Wang, Mien-Chie Hung, Zhongming Zhao, Ruiying Zhao, Chad D. Huff, Jingnan Shen, & Dung-Fang Lee. The spliceosome is an exploitable vulnerability in RB1-mutant cancers. *Nature Genetics* (Responsible for the TCGA data analysis part. Under review).

Published/Accepted
1. **Ruifeng Hu[#]**, Hao-Dong Xu[#], Peilin Jia, Zhongming Zhao. KinaseMD: Kinase mutations and drug responses in pan-cancer. *Nucleic Acids Research, gkaa945*.
2. Yulin Dai[#], **Ruifeng Hu[#]**, Peilin Jia, Zhongming Zhao. CSEA-DB: an omnibus for human complex trait and cell type associations, Nucleic Acids Research, gkaa1064 (<u>Co-first author</u>)
3. Guangsheng Pei[#], **Ruifeng Hu[#],** Yulin Dai, Astrid M Manuel, Zhongming Zhao, Peilin Jia. Predicting regulatory variants using a dense epigenomic mapped CNN model elucidated the molecular basis of trait-tissue associations. *Nucleic Acids Research, gkaa1137*(<u>Co-first author</u>)
4. Guangsheng Pei, **Ruifeng Hu**, Yulin Dai, Zhongming Zhao, Peilin Jia. Decoding whole-genome mutational signatures in 37 human pan-cancers by denoising sparse autoencoder neural network. *Oncogene*. 2020, 39, 5031–5041. (Responsible for VAE Model construction).
5. Hao-Dong Xu, **Ruifeng Hu**, Peilin Jia, Zhongming Zhao. 6mA-Finder: a novel method for identifying DNA N6-methyladenine sites in genome. *Bioinformatics* 36 (10), 3257-3259.
6. Yulin Dai, **Ruifeng Hu**, Zhongming Zhao, Peilin Jia. Diverse types of genomic evidence converge on alcohol use disorder risk genes. *Journal of Medical Genetics*. 2020 Mar 13.
7. Peilin Jia, Yulin Dai, **Ruifeng Hu**, Guangsheng Pei, Astrid Marilyn Manuel, and Zhongming Zhao. TSEA-DB: a trait–tissue association map for human complex traits and diseases. *Nucleic Acids Researh*. 2020, 48(D1):D1022-D1030.
8. Deyou Tang, Bingrui Li, Tianyi Xu, **Ruifeng Hu**, Daqiang Tan, Xiaofeng Song, Peilin Jia, and Zhongming Zhao. VISDB: a manually curated database of viral integration sites in the human genome. *Nucleic Acids Research*. 2020, 48(D1):D633-D641.
9. **Ruifeng Hu**, Guangsheng Pei, Peilin Jia, and Zhongming Zhao. Decoding regulatory structures and features from epigenomics profiles: a Roadmap-ENCODE Variational Auto-Encoder (RE-VAE) model. *Methods*. 2019, S1046-2023(19)30269-5.

10. **Ruifeng Hu**, Yulin Dai, Peilin Jia, and Zhongming Zhao. ANCO-GeneDB: annotations and comprehensive analysis of candidate genes for alcohol, nicotine, cocaine and opioid dependence. *Database* (Oxford). 2018, bay121.

11. Ke Yang, Yun Luo, Shan Lu, **Ruifeng Hu**, Yuyang Du, Ping Liao, Guibo Sun, and Xiaobo Sun. Salvianolic acid B and ginsenoside Re synergistically protect against Ox-LDL-induced endothelial apoptosis through the antioxidative and antiinflammatory mechanisms. *Frontiers in Pharmacology*. 2018, 9:662.

12. **Ruifeng Hu**, and Xiaobo Sun. Design of new traditional Chinese medicine herbal formulae for treatment of type 2 diabetes mellitus based on network pharmacology. *Chinese Journal of Natural Medicines*. 2017, 15(6):436-441.

13. Dan Yang, Junjie Shao, **Ruifeng Hu**, Haimei Chen, Ping Xie, and Chang Liu. Angiotensin II promotes the anticoagulant effects of rivaroxaban via angiotensin type 2 receptor signaling in mice. *Scientific Reports*. 2017, 7(1):369.

14. **Ruifeng Hu**, Guibo Sun, and Xiaobo Sun. LSCplus: a fast solution for improving long read accuracy by short read alignment. *BMC Bioinformatics*. 2016, 17(1):451.

15. **Ruifeng Hu**, and Xiaobo Sun. lncRNATargets: A platform for lncRNA target prediction based on nucleic acid thermodynamics. *Journal of Bioinformatics and Computational Biology*. 2016, 14(4):1650016.

16. **Ruifeng Hu**, Guomin Ren, Guibo Sun, and Xiaobo Sun. TarNet: an evidence-based database for natural medicine research. *PLoS One*. 2016,11(6):e0157222.

17. Leilei Zhang, Jing Jin, Lijing Zhang, **Ruifeng Hu**, Li Gao, Xiaowei Huo, Dongyu Liu et al. Quantitative analysis of differential protein expression in cervical carcinoma cells after zeylenone treatment by stable isotope labeling with amino acids in cell culture. *Journal of Proteomics*. 2015, 126:279-87.

18. Qidi Ai, Guibo Sun, Yun Luo, Xi Dong, **Ruifeng Hu**, Xiangbao Meng, and Xiaobo Sun. Ginsenoside Rb1 prevents hypoxia-reoxygenation-induced apoptosis in H9c2 cardiomyocytes via an estrogen receptor-dependent crosstalk among the Akt, JNK, and ERK 1/2 pathways using a label-free quantitative proteomics analysis. *RSC Advance*. 2015, 5:26346-26363.

19. **Ruifeng Hu**, Xiaoyan Xing, Guibo Sun and Xiaobo Sun. The prospect of using bioinformatics technology in the field of biological medicine in the era of big data. *Acta pharmaceutica Sinica*. 2014, 49(11):1512-9.

20. **Ruifeng Hu**, Xu Zeng, Weiwei Gao, Qian Wang, and Zhihua Liu. HRAS: a webserver for early warning of human health risk brought by aflatoxin. *Molecular Biology Reports*. 2013, 40(2):1181-7.

PRESENTATIONS/CONFERENCES

1. Zhongming Zhao, Peilin Jia, Yulin Dai, **Ruifeng Hu**, Guangsheng Pei, Astrid M Manuel (2020). A trait-tissue association landscape for human complex traits and diseases. AMIA 2020 Informatics Summit, Houston, TX, March 24, 2020 (selected for podium presentation)

2. **Ruifeng Hu**, Guangsheng Pei, Peilin Jia, and Zhongming Zhao (2019). Decoding regulatory structures and features from epigenomics profiles: a Roadmap-ENCODE Variational Auto-Encoder (RE-VAE) model. Presented at the 68th Annual Meeting of The American Society of Human Genetics, Houston, TX, USA, October 15-19, 2019, Abstract # 1515W [Reviewers' Choice Abstracts: The top 10% of poster abstracts]
3. Deyou Tang, Bingrui Li, Tianyi Xu, **Ruifeng Hu**, Zhongming Zhao (2019). VISDB: A comprehensive database for human disease-related virus integration sites. Presented at the 68th Annual Meeting of The American Society of Human Genetics, Houston, TX, USA, October 15-19, 2019, Abstract # 1567T
4. Peilin Jia P, Guangsheng Pei, **Ruifeng Hu**, Zhongming Zhao (2019). A variational autoencoder model for accurate imputation of drug response. Presented at the 68th Annual Meeting of The American Society of Human Genetics, Houston, TX, USA, October 15-19, 2019, Abstract # 834W
5. Consuelo Walss-Bass, **Ruifeng Hu**, Laura Stertz, Gabriel Fries, William Russell, Nicholas Crist, Glenn Sandberg, Zhongming Zhao, and Thomas Meyer. (2019) Proteomics of addiction: Postmortem brain analyses of cocaine and opioid disorder. Presented at the 68th Annual Meeting of The American Society of Human Genetics, Houston, TX, USA, October 15-19, 2019, Abstract # 1824W [Reviewers' Choice Abstracts: The top 10% of poster abstracts]
6. Consuelo Walss-Bass, **Ruifeng Hu**, Laura Stertz, Gabriel Fries, William Russell, Nicholas Crist, Glenn Sandberg, Zhongming Zhao, and Thomas Meyer (2019). Proteomics of addiction: Postmortem brain analyses of cocaine and opioid disorder. Presented at the XXVIIth World Congress of Psychiatric Genetics (WCPG), Los Angeles, CA, USA, October 26–31, 2019.
7. Consuelo Walss-Bass, Laura Stertz, Gabriel R Fries, **Ruifeng Hu**, William Russell, Glenn Sandberg, Zhongming Zhao, Rodrigo Grassi-Oliveira, Thomas Meyer (2019). Proteomics Analysis Suggests Demyelination as a Mechanism of Cocaine-Induced Neurotoxicity: Postmortem Brain Analyses of Cocaine Use Disorder. Neuropsychopharmacology 44 (Suppl 1), 511-512
8. **Ruifeng Hu,** Peilin Jia, Zhongming Zhao (2018). ANCO Genes: Annotations and comprehensive analysis of candidate genes for alcohol, nicotine, cocaine, and opiate (ANCO) dependence. Presented at the 67th Annual Meeting of The American Society of Human Genetics, San Deigo, CA, USA, October 16-20, 2018, Abstract #1397F.
9. China Information of Traditional Chinese Medicine Congress, Nov 22-23 2014, Beijing, China. **Ruifeng Hu,** Guomin Ren, Guibo Sun, Xiaobo Sun. Research on Chinese herbal compounds based on network pharmacology -- a case study of Chinese herbal compounds on the treatment of type II diabetes mellitus. Proceedings of Traditional Chinese Medicine Informatics Meeting in China (2014), page 129-134. [Second Prize of Excellent Paper]

AWARDS & SCHOLARSHIPS

2016.10	National Scholarship for Graduate Students, from Chinese Ministry of Education and Ministry of Finance (CAMS&PUMC) (the highest scholarship for graduate students)
2014 & 2015	First Prize Scholarship for Academic Excellence (CAMS&PUMC)
2010.10, 2011.10	National Encouragement scholarship
2009.10	National Scholarship for Undergraduate Students, from Chinese Ministry of Education and Ministry of Finance (NJFU) (the highest scholarship for undergraduate students)
2010, 2011	The First Prize Scholarship of Nanjing Forestry University

SELECTED HONORS

2014, 2015, 20016	Outstanding Graduate Students
2012.6	Outstanding Undergraduates
2011.10	Excellent Student Cadre, Advanced Class in Jiangsu Province, China
2010	Second Prize, Mathematical Contest in Modeling (Jiangsu, China)
2010, 2011	Second Prize, National Software Talent Contest (Finals), China
2009, 2010, 2011	Merit Student of Nanjing Forestry University, China (In three consecutive years)