

Curriculum Vitae

Brief Information

Name: Yiyong (Rambo) Zhao

Date of Birth: Jan 2, 1991

Nationality: China

Phone: +1 (857) 381-5442

Web: www.yiyongzhao.com

Address: 60 Fenwood Road, Boston, MA 02115, USA

Email: yzhao@bwh.harvard.edu or zhaoyiy@broadinstitute.org



Education

2016.06-2021.06: Fudan University, China -- Ph.D. in Bioinformatics (Integrated Master's and Ph.D. Program, QS World University Rankings 2024 = 50th)

2019.09-2020.10: Pennsylvania State University, USA (Major in Bioinformatics, Joint Ph.D. Program)

2012.06-2016.06: Northeast Forestry University, China; Bachelor of Agriculture (*summa cum laude*)

2013.06-2014.01: Kangwon National University, South Korea. (Joint Bachelor Program)

Work Experience

2023.05–present: Postdoctoral Research Fellow at Bioinformatics, Genomics and Bioinformatics Hub, Harvard Medical School and Brigham & Women's Hospital, Boston, MA, USA. (Project: Expression Quantification of Transposable Elements and Association Analysis in Neurodegenerative Diseases by Integrating Multi-omics Data)

2023.05–present: Affiliate Faculty Member of Broad Institute of Harvard and MIT, Cambridge, MA, USA.

2022.10-2023.04: Professor at Department of crop genetic breeding, Guizhou University, China.

2021.07-2022.08: Drug Farm Co. Ltd (<https://drug-farm.com>), Shanghai, China as Bioinformatics & AI Scientist.

Major Courses

Mathematical Methods of Statistics, Biostatistics: Theories and Applications; Linear Statistical Analysis; Genomics and Bioinformatics; Statistical Analysis and Data Mining of Omics Data; Methods and Applications of Evolutionary Genomics, Analysis and Applications of Microarray and

Next-generation Sequencing Data, Analysis of Categorical Data and Survival Data, Text Mining, The Mathematical Theory and Algorithms for Machine Learning and Pattern Recognition, Big Data-Oriented Machine Learning Theory and Algorithms and etc.

Research Interests

Computational biology, Biological big data mining; Genomics, Revealing the unknown functions of the human genome and their roles in neurological diseases by applying and developing computational methods to integrate multi-dimensional multi-source data, Phylogenomic and comparative genomics; Bioinformatics software development; Genome assembly, Genome evolution such as whole genome duplications; Hybridization detection; Gene evolution and functional genomics.

Publications (#Co-first author, *Corresponding author)

Published/Accepted (H-index: 9; Citations: 649):

[1] **Yiyong Zhao**, Rong Zhang, Kai-Wen Jiang, Ji Qi, Yi Hu, Jing Guo, Renbin Zhu, Taikui Zhang, Ashley N. Egan, Ting-Shuang Yi, Chien-Hsun Huang, and Hong Ma, Nuclear phylotranscriptomics and phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in Fabaceae. *Molecular Plant*, 2021. 14(5): p. 748-773. (Cover story & Featured article, **IF₂₀₂₁=21.949**, reported in dozens of domestic and international mainstream media such as Guangming Daily, People's Daily, China News, Science Daily, Phy.org, etc., My contribution: for the cover article, based on the transcriptome and genome, I constructed the largest phylogenetic relationship of Leguminosae to date. Phylogenetic relationships were resolved for five subfamilies, about 30 genome-wide replication events were identified, the evolutionary history of the nitrogen-fixing gene family was explored, and the important role of a new gene in nitrogen fixation was proposed, providing an important basis for germplasm conservation and breeding in the legume family. (The full analyses presented in this article were conducted by me, drafted the original manuscript and made rebuttal.)

#[2] Liangsheng Zhang[#], Fei Chen[#], Xingtian Zhang[#], Zhen Li[#], **Yiyong Zhao**[#], Rolf Lohaus[#], Xiaojun Chang[#], Wei Dong, Simon Y. W. Ho, Xing Liu, Aixia Song, Junhao Chen, Wenlei Guo, Zhengjia Wang, Yingyu Zhuang, Haifeng Wang, Xuequn Chen, Juan Hu, Yanhui Liu, Yuan Qin, Kai Wang, Shanshan Dong, Yang Liu, Shouzhou Zhang, Xianxian Yu, Qian Wu, Liangsheng Wang, Xueqing Yan, Yuannian Jiao, Hongzhi Kong, Xiaofan Zhou, Cuiwei Yu, Yuchu Chen, Fan Li, Jihua Wang, Wei Chen, Xinlu Chen, Qidong Jia, Chi Zhang, Yifan Jiang, Wanbo Zhang, Guanhua Liu, Jianyu Fu, Feng Chen, Hong Ma, Yves Van de Peer, and Haibao Tang, The water lily genome and the early evolution of flowering plants. *Nature*, 2020. 577(7788): p. 79-84. (**IF₂₀₂₀=49.962, ESI highly cited paper**), (my contribution involved in the assembly, annotation, and correction of transcriptome data of the first aquatic early angiosperm genome. Based on 115 transcriptome and genome-wide data, a highly supported phylogenetic tree of flowering plants was constructed, and molecular clock analysis revealed that early angiosperms originated in the early Cretaceous, while highly supporting the phylogenetic relationships within the Water Lily family. Low-copy nuclear genes from the whole genome confirmed that saprophytic camphor was the earliest differentiated angiosperm. Meanwhile, using a comparative genomics approach, the Blue Star water lily was

identified as a potential hybrid parent of two cultivars, laying the foundation for genomics in molecular cultivation breeding. (The related analyses of Phylogenomic relationships and divergence time estimations of angiosperms in Figure 1 were conducted by me and drafted the corresponding part manuscript and made rebuttal.)

*[3] Lin Cheng, Qunwei Han, Fei Chen, Mengge Li, Tiago Santana Balbuena, and **Yiyong Zhao**^{*}, Phylogenomics as an effective approach to untangle cross-species hybridization event: A case study in the family Nymphaeaceae. **Frontiers in Genetics** (section in computational genomics), 2022. 13. (IF₂₀₂₂=3.7) (My contributions in conducting nearly all analyses, drafted the original manuscript, and handled revisions and rebuttals.)

*[4] Lin Cheng[#], Mengge Li[#], Yachao Wang[#], Qunwei Han, Yanlin Hao, Zhen Qiao, Wei Zhang, Lin Qiu, Andong Gong, Zhihan Zhang, Tao Li, Shanshan Luo, Linshuang Tang, Daliang Liu, Hao Yin, Song Lu, Tiago Santana Balbuena, and **Yiyong Zhao**^{*}, Transcriptome-based variations effectively untangling the intraspecific relationships and selection signals in Xinyang Maojian tea population. **Frontiers in Plant Science**, 2023. 14. (IF₂₀₂₂=5.6) (My contributions in conceptualization and supervision and handled revisions and rebuttals.)

[#][5] Angiosperm-wide analysis of fruit and ovary evolution aided by a new nuclear phylogeny supports association of the same ovary type with both dry and fleshy fruits. Yezi Xiang[#], Taikui Zhang[#], **Yiyong Zhao**[#], Hongjin Dong, Hongyi Chen, Yi Hu, Ji Qi, Chien-Hsun Huang, Jun Xiang^{*}, and Hong Ma^{*}. **Journal of Integrative Plant Biology**, 2024, 00: 1-24. (IF₂₀₂₂=11.4) (My contributions including examined available datasets and chose the high-quality datasets for 423 angiosperm species, representing all major angiosperm lineages and nearly all orders. Due to large scale of hundreds of angiosperm species and many thousands of genes, the molecular evolutionary analyses took a lot of steps and a lot of time. I was able to generate the appropriate gene phylogenies and summarized the results for the purpose of reporting gene duplications at key nodes in the angiosperm phylogeny as reported in this paper.)

[6] Daliang Liu, Zhihan Zhang, Yanlin Hao, Mengge Li, Houlin Yu, Xingruo Zhang, Haoyang Mi, Lin Cheng^{}, **Yiyong Zhao**^{*}. Decoding the complete organelle genomic architecture of *Stewartia sinensis*: An early-diverging species in Theaceae with phylogenetic implications. 25, 114 (2024). **BMC Genomics** (IF₂₀₂₂=4.4) (My contribution in conceptualization and supervision and handled revisions and rebuttals.)

*[7] Lin Cheng[#], Mengge Li[#], Qunwei Han, Zhen Qiao, Yanlin Hao, Tiago Santana Balbuena, and **Yiyong Zhao**^{*}, Phylogenomics resolves the phylogeny of Theaceae by using low-copy and multi-copy nuclear gene makers and uncovers a fast radiation event contributing to tea plants diversity. **Biology**, 2022. 11(7): p. 1007. (IF₂₀₂₂=4.2) (My contribution in conceptualization and supervision and handled revisions and rebuttals.)

[#][8] Yifan Jiang^{#, *}, Fei Chen[#], Aixia Song[#], **Yiyong Zhao**[#], Xinlu Chen, Yuan Gao, Guo Wei, Wanbo Zhang, Yaqin Guan, Jianyu Fu, Junhao Chen, Wenlei Guo, Zhengjia Wang, Tobias G. Köllner, Zhineng Li, Ning Liu, Shunzhao Sui, Liangsheng Zhang^{*}, and Feng Chen^{*}, The genome assembly of *Chimonanthus praecox* var. *concolor* and comparative genomic analysis highlight the genetic basis underlying conserved and variable floral traits of wintersweet. **Industrial Crops and Products**, 2023. 206: p. 117603. (IF₂₀₂₂=5.9) (The related analyses of phylogenomic relationships and divergence time estimations of magnoliids in Figure 4 were conducted by me and drafted the

corresponding part manuscript and made rebuttal.)

[9] Lin Cheng[#], Yanlin Hao[#], Qunwei Han, Zhen Qiao, Mengge Li, Daliang Liu, Hao Yin, Tao Li, Wen Long, Shanshan Luo, Ya Gao, Zhihan Zhang, Houlin Yu, Xinhao Sun, **Yiyong Zhao**^{*}. The origin and evolution of orphan genes: a case study in tea plant family. *bioRxiv* 2024.02.01.578514; doi: <https://doi.org/10.1101/2024.02.01.578514> (My contribution in conceptualization and supervision and handled revisions and rebuttals.)

*[10] Dandan Li[#], Mingjin Huang[#], Qunwei Han, Dachang Wanag, Kaijie Li, Qiuyue Yang, Ronghui Gu, Guangchong Zhou, Songtao He, Houlin Yu, Lin Cheng, **Yiyong Zhao**^{*}. A high-quality chromosomal-level reference genome of *Dendrobium nobile* L. provides new insights into the biosynthesis and accumulation of picrotoxane-type sesquiterpenoid alkaloids, 2024. <https://doi.org/10.1016/j.indcrop.2024.118243>. *Industrial Crops and Products* (IF₂₀₂₂=5.9) (My contribution in Conceptualization, Data curation, Investigation, Writing/review/editing & handled revisions and rebuttals.)

[11] Lin Zhang, Xinxin Zhu, **Yiyong Zhao**, Jing Guo, Taikui Zhang, Weichen Huang, Jie Huang, Yi Hu, Chien-Hsun Huang^{*}, and Hong Ma^{*}, Phylotranscriptomics resolves the phylogeny of Pooideae and uncovers factors for their adaptive evolution. *Molecular Biology and Evolution*, 2022. 39(2). (IF= 16.24) (My contributions including taxon samplings and sequencing, assistance in Lin's data analyses)

[12] Jing Guo, Weibin Xu, Yi Hu, Jie Huang, **Yiyong Zhao**, Lin Zhang, Chien-Hsun Huang, and Hong Ma, Phylotranscriptomics in Cucurbitaceae reveal multiple whole-genome duplications and key morphological and molecular innovations. *Molecular Plant*, 2020. 13: p. 1-17. (IF=13.164) (My contributions including taxon samplings and sequencing, assistance in Jing's data analyses)

[13] Chengpeng He[#], Zhiyu Chen[#], **Yiyong Zhao**, Yue Yu, Hongkuan Wang, Cong Wang, Gregory P. Copenhaver, Ji Qi^{*}, and Yingxiang Wang^{*}, Histone demethylase IBM1-mediated meiocyte gene expression ensures meiotic chromosome synapsis and recombination. *PLOS Genetics*, 2022. 18(2): p. e1010041. (IF₂₀₂₂=5.917) (My contributions including all the gene expression quantification and differential gene expressions analyses from RNA-seq datasets)

[14] Jie Huang, Weibin Xu, Junwen Zhai, Yi Hu, Jing Guo, Caifei Zhang, **Yiyong Zhao**, Lin Zhang, Christopher Martine, Hong Ma, and Chien-Hsun Huang, Nuclear phylogeny and insights into whole-genome duplications and reproductive development of Solanaceae plants. *Plant Communications*, 2023. 4(4): p. 100595. (IF₂₀₂₂=10.5) (My contributions including taxon samplings and sequencing, assistance in Jie's data analyses)

[15] Yinghai Liang, Shanshan Wang, Chenhui Zhao, Xinwei Ma, **Yiyong Zhao**, Jing Shao, Yuebo Li, Honglian Li, Hongwei Song, Hong Ma, Hao Li, Bingbing Zhang, and Liangsheng Zhang, Transcriptional regulation of bark freezing tolerance in apple (*Malus domestica* Borkh.). *Horticulture Research*, 2020. 7(1): p. 205. (IF₂₀₂₀=6.072) (My contributions including the gene expression quantification and differential gene expressions analyses for RNA-seq data)

[16] Zhuang Meng, Jinlei Han, Yujing Lin, **Yiyong Zhao**, Qingfang Lin, Xiaokai Ma, Jianping Wang, Muqing Zhang, Liangsheng Zhang, Qinghui Yang, and Kai Wang, Characterization of a *Saccharum spontaneum* with a basic chromosome number of x = 10 provides new insights on genome evolution in genus *Saccharum*. *Theoretical and Applied Genetics*, 2020. 133(1): p. 187-

199. (IF₂₀₂₀=5.699) (My contributions including Phylogeny and divergence time estimation of Poaceae as in Figure 5.)

[17] Weichen Huang, Lin Zhang, J. Travis Columbus, Yi Hu, **Yiyong Zhao**, Lin Tang, Zhenhua Guo, Wenli Chen, Michael McKain, Madelaine Bartlett, Chien-Hsun Huang, De-Zhu Li, Song Ge, and Hong Ma, A well-supported nuclear phylogeny of Poaceae and implications for the evolution of C₄ photosynthesis. *Molecular Plant*, 2022. 15(4): p. 755-777. (IF₂₀₂₂= 27.5) (My contributions including taxon samplings and sequencing, assistance in Weichen's data analyses)

[18] Fanjuan Meng, Li Liu, Mu Peng, Zhongkui Wang, Chao Wang, and **Yiyong Zhao**, Genetic diversity and population structure analysis in wild strawberry (*Fragaria nubicola* L.) from Motuo in Tibet Plateau based on simple sequence repeats (SSRs). *Biochemical Systematics and Ecology*, 2015. 63: p. 113-118. (IF₂₀₁₅=1.381) (My contributions including taxon samplings and sequencing, assistance in Mu's data analyses)

[19] Ye Zheng, **Yiyong Zhao**, Lihua Wu, Juan Lin. Study on copy number variation and differentiated expression pattern of *CaCA* family genes in *Capsella*. 2020. <http://www.jns.fudan.edu.cn/#/digest?ArticleID=1178>, *Chinese Journal of Fudan University (Natural Science)* [J], 59, No.1:1-31 (My contributions including the gene expression quantification and differential gene expressions analyses for RNA-seq data)

[20] **Yiyong Zhao^{#,*}, Daliang Liu[#], Zhihan Zhang[#], Song Lu, Lin Cheng, Houlin Yu, Hao Yin, Tao Li, Yabing Hou^{#,*}. Comparative transcriptomic and proteomic analysis of hypoxia response in wild and cultivated tomato roots. 2023. *BMC Genomics* (IF₂₀₂₂=4.4, Under review)

[21] Mingjin Huang^{#,*}, Shanshan Luo[#], Linshuang Tang[#], Hao Yin[#], Daliang Liu, Zhipeng Li, Qiyu Chen, Yinjie Jiao, Mengge Li, Zhihan Zhang, Dachang Wang, Lin Cheng, Cheng Li, Houlin Yu, **Yiyong Zhao^{#,*} The Integration of Pan-genome and Multi-omics Analyses Provides Insights into the Evolution of the Mycoheterotrophic Orchid *Gastrodia elata*. 2024. *Communications Biology* (IF₂₀₂₂=4.4, Under review) (My contribution in conceptualization, investigation, supervision, drafted the initial manuscript, will handle revisions and rebuttals)

[22] Daliang Liu, Hao Yin, Tao Li, Liang Wang, Quanzhi Zhao^{}, **Yiyong Zhao**^{*}. The origins of cultivated rice and genomic signature of heterosis for yield traits in super-hybrid rice. 2024. *Theoretical and Applied Genetics* (IF₂₀₂₂=4.4, Under review) (My contribution in conceptualization and supervision, assistance in Daliang's data analyses, drafted the initial manuscript, will handle revisions and rebuttals.)

*[23] The nearly complete genome of an early-diverging species *Stewartia gemmata* and phylogenomic studies provide insights into new gene evolution in Theaceae. Lin Cheng[#], Yanlin Hao[#], Zhen Qiao, Qunwei Han, Daliang Liu, Hao Yin, Tao Li, Wen Long, Shanshan Luo, Ya Gao, Hao Li^{*}, **Yiyong Zhao**^{*}. *Communications Biology* (IF₂₀₂₂=4.4, Under review) (My contribution including collected samples and sequencing, conceptualization and supervision, assistance in Yanlin's data analyses, drafted the initial manuscript, will handle revisions and rebuttals)

#[24] **Yiyong Zhao**, Wenyu Kuo, Jie Huang, Miao Sun, Yi Hu, Jing Guo, Pamela S. Soltis, Douglas E. Soltis, Hong Ma^{*} and Chien-Hsun Huang^{*}, Nuclear phylogenomics provide evidence to clarify key morphological evolution and whole genome duplication across rosids. 2024. *Journal of Integrative Plant Biology* (IF₂₀₂₂=11.4, Under review) (My contribution including collected

samples and sequencing, transcriptome assembly, phylogenomic analyses, ancestral morphological state analysis, and drafted the initial manuscript, will handle revisions and rebuttals)

Programming Skills

Proficient in Linux, Python, R and other programming languages for big data mining.

Developed Software/Web

PhyloTracer: A User-Friendly Toolkit for Gene Tree Rooting Gene Duplication, Expansion and Contraction Identification. Ortholog Retrieval, Phylogenetic Noise Elimination, Species Hybridization Detection, and Visualization.

Python version: <https://github.com/YiyongZhao/PhyloTracer>

Web version: <http://www.bioinfordocker.com/PhyloTracer> (Under construction)

FabGenomeGalaxy: An Integrative Platform for Fabaceae Comparative Genomics and Gene Family Analysis.

Database derived commutating platform: <http://www.bioinfordocker.com/FabGenomeGalaxy> (Under construction)

Editorial Activities

- Ad-hoc Reviewer –

Frontiers in Bioinformatics, Frontiers in Genetics, Horticulture Research, Frontiers in Plant Science, Frontiers in Molecular Biosciences, Journal of Systematics and Evolution, BMC Ecology and Evolution, Frontiers in Endocrinology, Open Medicine and Plant Genetic Resources.

- Editorial Roles –

2023 - present: As a member of Reviewer Board in MDPI academic press (<https://www.mdpi.com>)

2023 - present: Frontiers in Plant Science specialty section of Functional and Applied Plant Genomics. (Academic Editor)

2023 - present: Frontiers in Ecology and Evolution (Academic Editor)

2023 - present: Innovation (Academic Editor)

2022 - present: Tropical Plants (Academic Editor)

Faculty Academic Appointments

2023-present: Deputy Director and Technical & Academic Lead, Key Laboratory of Molecular Breeding of Grain and Oil Crops of Guizhou Province, College of Agriculture, Guizhou

University, Guiyang 550025, China. (Responsibilities: Gene Mining for Significant Agronomic Traits in Crops).

Conferences or Training

[1] 2023.08.22-08.26: Eukaryotic mRNA Processing Conference, Topics: Mechanisms of RNA Splicing, Alternative Splicing: Mechanisms, Networks, and Biological Consequences, 3' End Processing, RNA Modifications, RNA Structure in Cells, Co-transcriptional RNA Processing: Dynamics of co-transcriptional RNA processing, RNA Turnover and Quality Control, Viruses and RNA Processing, RNA Processing in Disease, RNA-Protein Interactions/RNA Complexes; Location: Cold Spring Harbor Laboratory, NY, USA.

[2] 2020.07.27-07.31: American Society of Plant Biologists: The Plant Biology 2020 Worldwide Summit, Location: USA.

[3] 2019.03.30: AI Marketing-MobTech Precision Marketing, Location: 2F, Yuyu Building, Building 2, Huaxin Center, No. 711 Yishan Road, Xuhui, Shanghai.

[4] 2018.12.13: Workshop on the Intersection of Artificial Intelligence and Biomedical Big Data, Location: Guanghua Building, Fudan University.

[5] 2018.12.08-12.09: The 11th China-R Conference (Shanghai) and East China Data Science Conference, Location: East China Normal University.

[6] 2018.12.01: Big Data and Artificial Intelligence 2018 Ph.D. Forum, Location: Zhangjiang Campus of Fudan University.

[7] 2018.11.10: International Conference on the 10th Anniversary of the Institute of Plant Sciences, Location: Jiangwan Campus of Fudan University.

[8] 2018.10.29: Data Castle's Competition: Car Destination Intelligence Prediction Contest.

[9] 2018.10.26: Data Man Open Class: Big Data Business Application Analysis and Future Direction Outlook, Location: Room H6412, Fudan University.

[10] 2018.10.21: The 8th Python Developers Conference PyConChina2018, Location: Shanghai Putuo District.

[11] 2018.09.19: World Artificial Intelligence Conference, Sub-forum: Heterogeneous Computing, Standards Building, Industry Empowerment.

[12] 2018.09.14: The 15th China Graduate Student Mathematical Modeling Competition.

[13] 2018.09.11-09.12: International Summit on Computer Vision Technology and Applications (as a Volunteer), Location: Shanghai Putuo District.

[14] 2018.06.24: Shanghai Workshop on Neuroimaging and Applications, Location: Zibin Building, Fudan University.

[15] 2018.04.20-04.22: The 5th National Conference on Computational Biology and Bioinformatics and the 2nd Phoenix Genome Informatics Forum.

[16] 2018.04.09: Data Castle's Competition: Artificial Intelligence-based Drug Molecular Screening Competition.

[17] 2018.04.29: Big History of Artificial Intelligence by Prof. Tiejun Huang, Location: Guanghua Building, Fudan University.

[18] 2018.02.24-02.27: The 2nd Symposium of the International MAQC Society for Quality Control of Omics Data, Location: Fudan University.

[19] 2018.01.31-02.04: Workshop on Molecular Paleobiology: Phylogeny and Divergence Time Estimation, Organizers: Yin Zongjun (NIGPAS) and Philip Donoghue (University of Bristol, UK), Location: Nanjing Institute of Paleontology.

[20] 2017.12.20-12.21: The 2nd Conference on Artificial Intelligence for Precision Biomedical Informatics, Location: Wu Wenzheng Lecture Hall, Fudan Guanghua Building.

[21] 2017.12.02-12.03: The 10th China-R Conference (Shanghai) and East China Data Science Conference, Location: East China Normal University.

[22] 2017.12: Fudan Science and Technology Innovation Forum, Location: Shanghai Wujiaochang.

[23] 2017.11.01-11.03: 2017 IUFRO Genetics and Genomics of Fagaceae (as a volunteer).

[24] 2017.07.23-07.29: XIX International Botanical Congress, Location: Shenzhen.

[25] 2017.07.17-07.21: Microbial Genomics Training Course by Shanghai Majorbio Bio-pharm Technology Co., Ltd.

[26] 2017.07.02-07.07: 2017 Fudan University Big Data Institute Summer Course, Location: Fudan Guanghua Building.

[27] 2017.06.25-07.02: The 13th Fudan University Summer Advanced Workshop on Ecology, Location: Ecology Building of Fudan University.

[28] 2017.06.26-06.27: 2017 Pujiang Genetic and Molecular Diagnostics Industry-Academia Research Forum.

[29] 2017.06.16-06.18: Bioinformatics and Intelligent Information Processing Academic Conference, Location: Minhang Campus, Shanghai Jiao Tong University.

Foundation

Participated in two General Program of National Natural Science Foundation of China:

[1] National Natural Science Foundation of China, General Project, 31970224, Reconstruction of Cucurbitaceae Systematics Based on Nuclear Genes and Related Evolutionary Biology Studies, from 2020-01-01 to 2023-12-31, 580,000 RMB, Ongoing, as Participant.

[2] National Natural Science Foundation of China, General Project, 31770242, Constructing a Comprehensive Systematic Evolutionary Relationship of the Brassicaceae and Analyzing the Evolutionary Pathway of Flower and Fruit Development and Stress-Resistant Genes in Different Lineages, from 2018-01-01 to 2019-12-31, 250,000 RMB, Completed, as Participant.

Selected Honors and Awards

[1] 2016-2021 The First Prize Graduate Academic Scholarship of Fudan University

- [2] 2016 Northeast Forestry University "Outstanding Graduate" Award
- [3] 2015 Third Prize in the Second Plant Classification Competition Summer Camp jointly organized by the Chinese Academy of Sciences Institute of Botany, Peking University, and Beijing Forestry University
- [4] 2013-2014 Northeast Forestry University "Three Good Student" Award
- [5] 2013 National Encouragement Scholarship
- [6] 2012-2013 Northeast Forestry University First and Second Level Scholarships
- [7] 2012 Northeast Forestry University "Hand-in-Hand" Outstanding Academic Scholarship

Key Contributions and Achievements at Drug Farm (2021-2022)

Role: Bioinformatics & Artificial Intelligence Scientist at Drug Discovery and Indication Predictions

[1] Indication Predictions:

- a) Spearheaded the examination and determination of drug indications, focusing on ailments such as Kawasaki disease, systemic lupus erythematosus, IgA nephropathy, diabetic nephropathy, atherosclerosis, acute myocardial infarction, ischemic stroke, idiopathic pulmonary fibrosis, PBI nephropathy (with UUO, adenine, Cisplatin treatment), type 2 nephropathy due to diabetes mellitus, vascular shear stress and atherosclerosis, acute myeloid leukemia, and myelodysplastic syndrome.
- b) Employed natural language processing for swift data examination, accelerating the identification of potential indications for drugs.

[2] Drug Target Discovery:

- a) Collaboratively introduced IDinvivo+, an innovative tool designed for the prompt discovery of novel drug targets. This tool harnessed the power of graph neural networks founded on natural language processing, amalgamating data from renowned databases such as PPI, GO, KEGG, MeSH, ChemBL, GEO, and more.
- b) Conducted gene enrichment analyses rooted in curated databases, utilizing the most recent software and data sources. This ensured our enrichment assessments (pertaining to GO, KEGG, Diseases) were consistently in sync with the latest industry data.
- c) Orchestrated a comprehensive Meta-Analysis using various datasets, combined with the P-value method, to assess the merits of distinct algorithms like Empirical Fisher's methods, Stouffer's Methods, among others. This led to the development of "Meta-analysis_V9_ZYY", an advanced automated meta-analysis program that integrates the Network Analyst for thorough data assessment.
- d) Furthered our co-expression analyses by grasping different clustering algorithms, ultimately adopting the Fuzzy C-Means Clustering and K-means techniques to decipher the relationship between OSS shear force and atherosclerosis.

[3] Technical Mastery & Infrastructure Deployment:

- a) Oversaw the deployment of a supercomputing cluster server optimized for GPU

functionality, ensuring swift and seamless data computations.

- b) Spearheaded web-crawling initiatives, gleaning valuable data from prominent sources such as NCBI, Malacards diseases, Pharmado, and the likes.
- c) Demonstrated adeptness in deploying and leveraging a spectrum of biomedical databases including HGNC, Uniprot, DO disease, Europe_PMC, and more.
- d) Cultivated expertise in crucial bioinformatics tools such as MetaMap, NetworkX, PubMed_parser, Bio.Medline, Bio.Entrez, Rank-In, and Network Analyst.
- e) Achieved proficiency in deciphering single-cell sequencing datasets, diverse microarray, and transcriptome data processing.

During my tenure at Drug Farm, I immersed myself in a blend of cutting-edge methodologies and in-depth bioinformatics studies, continuously pioneering advancements in drug discovery.

Volunteer Experience

[1] June 2014 - August 2014: Participant in AIESEC Global Community Development Program

- a) Actively engaged in the "Heart to Heart" project in collaboration with CFPA, aimed at poverty alleviation and educational support in poor area Yichun, Wuyiling, China.
- b) Responsible for cultural exchange initiatives, enhancing community engagement through teaching experiences and proficiency in public speaking.
- c) Completed volunteer service that fostered teamwork skills and provided in-depth understanding of educational needs in underprivileged areas.

[2] Regular Blood Donor, Red Cross: Donated over 1000ml of blood, aiding critical medical needs and contributing to life-saving supplies.