CURRICULUM VITAE

Chang Liu

E-mail: chang.liu.cl2549@yale.edu

Address: 10 Sachem St, Rm.226, New Haven, CT 06520

Phone: (+1) 203-627-6465

Education Background

Sept. 2022 -	**Postdoc**, Department of Anthropology, Yale University, New Haven, U.S.
Sept. 2015 - May 2022	**Ph.D. Candidate**, Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai, China
Mar. 2019 - May 2019	**Visiting student**, Computational Biology, Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
Sept. 2018	**Visiting student**, Computational Biology, School of Computing Sciences, University of East Anglia, Norwich, UK.
Sept. 2011 - Jul. 2015	**B.S.**, Fundamental Mathematics, School of Mathematics, Shandong University, Jinan, China.
Sept. 2014 - May 2015	**Exchange Student**, Computational Biology, Institute of Systems Science, Academy of Mathematics and Systems Science, CAS, Beijing, China.

Honors/Awards

Nov. 2017	China National Scholarship
May 2019, May 2017	Merit Student of CAS

Teaching Activity

Mar. 2020 - Jun.2020	Biostatistics, Teaching Assistant, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences

Sept. 2020 - Jan.2021   Principles of Computational Biology, Teaching Assistant, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences

Research Interests

- **Admixture history inference**

  Developing methods for admixture history inference. Tracing ancestral origins of admixed populations and exploring characterization of population genetics of ancestral populations.

- **Local ancestry inference and archaic introgression detection**

  Developing methods for inferring local ancestry of admixed genomes and seeking segments of archaic introgression on modern human genomes.

Research Experiences

- **Local ancestry inference and archaic introgression detection**
  - We developed ***ArchaicSeeker2.0*** under the Hidden Markov Model to identify introgressed hominin sequences. We further inferred the ancestry of each introgressed sequences and the divergence time by a likelihood approach. A novel statistic, archaic ancestry-inference, has been proposed to quantify the similarity of archaic introgression distribution across genomes among two populations. We discerned two waves of introgression in present-day Eurasian populations and an ancient Siberian individual.
  - **Collaborator: Dr. Kai Yuan,** Massachusetts General Hospital and Broad Institute, Boston, U.S.

    Dr. Xumin Ni, Beijing Jiaotong University, Beijing, China

- **Admixture history inference**
  - We developed ***MultiWaver2.0*** modeling the length distribution of ancestral tracks by a mixed exponential distribution and estimating admixture proportions and dates by EM algorithm. It can distinguish continuous and multiple-wave discrete admixture by Bayesian Information Criterions. The method has been used for exploring admixture histories of ethnic minorities of northwest China and introgression histories of archaic hominins.

- I'm developing a novel method which can infer both admixture and divergence date of each admixture wave in multiple-wave admixture and further construct a more complete history model by a novel framework. The admixture wave from which each ancestral segment originated can be inferred as well.
- **Collaborator: Dr. Xumin Ni,** Beijing Jiaotong University, Beijing, China
  Dr. Kai Yuan, Massachusetts General Hospital and Broad Institute, Boston, U.S.

- **Simulation software development**
  - We developed *AdmixSim2*, a flexible forward-time simulation tool, to simulate real sequence data in a series of complex admixture processes with various evolutionary driving forces including de novo mutation, drift, recombination, and natural selection. It can also be applied in admixture scenarios for sex chromosomes and non-human species.
  - **Collaborator: Rui Zhang,** CAS-MPG Partner Institute for Computational Biology, Shanghai, China

- **Ancestry informative markers (AIMs) selection**
  - I developed an approach to optimize AIMs panels by a splicing algorithm. The optimized panels performed better than general panels with top markers or progressively screened panels. I'm further developing a tool that integrates a screening module, an optimizing module and a classifying module. It has been partially used for analyzing individuals' ancestries as a function of *PGG.Han* dataset using AIMs screened by the tool.

- **PRDM9 genetic diversity characterization**
  - I have explored to characterize PRDM9 genetic diversity in human populations by short sequence reads. I simulated PRDM9-fingerprints of known PRDM9 genotype and leveraged various machine learning approaches to classify the PRDM9-fingerprints of an individual into simulated categories.

Publications

[1] Kai Yuan#, Xumin Ni#, **<u>Chang Liu</u>**#, Yuwen Pan#, Lian Deng#, Rui Zhang#, Yang Gao, Xueling Ge, Jiaojiao Liu, Xixian Ma, Haiyi Lou, Taoyang Wu, Shuhua Xu*. 2021. Refining models of archaic admixture in Eurasia with *ArchaicSeeker 2.0*. **Nature Communications** 12(1): 6232.

> **Contribution**: I was independently responsible for the development, evaluation and analysis of a novel statistic, and jointly responsible for mathematical modeling, statistical analysis and algorithm development.

[2] Rui Zhang#, **Chang Liu**#, Kai Yuan, Xumin Ni, Yuwen Pan and Shuhua Xu*. 2021. *AdmixSim 2*: a forward-time simulator for modeling complex population admixture. **BMC Bioinformatics** 22(1): 506.

> **Contribution**: I was independently responsible for mathematical modeling and software development in Python.

[3] Xumin Ni, Kai Yuan, **Chang Liu**, Qidi Feng, Lei Tian, Zhiming Ma*, Shuhua Xu*. 2018. *MultiWaver 2.0*: modeling discrete and continuous gene flow to reconstruct complex population admixtures. **European Journal of Human Genetics**. 27:133-139.

> **Contribution**: I was jointly responsible for mathematical modeling and evaluation in method development.

[4] Qidi Feng#, Yan Lu#, Xumin Ni#, Kai Yuan#, Yajun Yang#, Xiong Yang, **Chang Liu**, Haiyi Lou, Zhilin Ning, Yuchen Wang, Dongsheng Lu, Chao Zhang, Ying Zhou, Meng Shi, Lei Tian, Xiaoji Wang, Xi Zhang, Jing Li, Asifullah Khan, Yaqun Guan, Kun Tang*, Sijia Wang*, Shuhua Xu*. 2017. Genetic history of Xinjiang's Uyghurs suggests Bronze Age multiple-way contacts in Eurasia. **Molecular Biology and Evolution**. 34(10):2572–2582.

[5] Yang Gao#, Chao Zhang#, Liyun Yuan#, Yunchao Ling, Xiaoji Wang, **Chang Liu,** Yuwen Pan, Xiaoxi Zhang, Xixian Ma, Yuchen Wang, Yan Lu, Kai Yuan, Wei Ye, Jiaqiang Qian, Huidan Chang, Ruifang Cao, Xiao Yang, Ling Ma, Yuanhu Ju, Long Dai, Yuanyuan Tang, The Han100k Initiative, Guoqing Zhang*, Shuhua Xu*. 2020. PGG.Han: the Han Chinese genome database and analysis platform. **Nucleic Acids Research** 48(D1): D971-D976.

[6] Chao Zhang#, Yang Gao#, Zhilin Ning#, Yan Lu#, Xiaoxi Zhang, Jiaojiao Liu, Bo Xie, Zhe Xue, Xiaoji Wang, Kai Yuan, Xueling Ge, Yuwen Pan, **Chang Liu**, Lei Tian, Yuchen Wang, Dongsheng Lu, Boon-Peng Hoh, Shuhua Xu*. 2019. PGG.SNV: understanding the evolutionary and medical implications of human single nucleotide variations in diverse populations. **Genome Biology** 20(1): 215.

[7] Lian Deng#, Haiyi Lou#, Xiaoxi Zhang#, Bhooma Thiruvahindrapuram, Dongsheng Lu, Christian R Marshall, **Chang Liu**, Bo Xie, Wanxing Xu, Lai-Ping Wong, Chee-Wei Yew, Aghakhanian Farhang, Rick Twee-Hee Ong, Mohammad

Zahirul Hoque, Abdul Rahman Thuhairah, Bhak Jong, Maude E Phipps, Stephen W Scherer, Yik-Ying Teo, Subbiah Vijay Kumar*, Boon-Peng Hoh*, Shuhua Xu*. 2019. Analysis of five deep-sequenced trio-genomes of the Peninsular Malaysia Orang Asli and North Borneo populations. **BMC Genomics** 20(1): 842.

[8] Kai Yuan#, Ying Zhou#, Xumin Ni#, Yuchen Wang, <u>**Chang Liu**</u> and Shuhua Xu*. 2017. Models, methods and tools for ancestry inference and admixture analysis. **Quantitative Biology**. 5(3):236-250.

[9] Haiyi Lou#*, Yang Gao#, Bo Xie#, Yimin Wang#, Haikun Zhang, Miao Shi, Sen Ma, Xiaoxi Zhang, <u>**Chang Liu**</u>, Shuhua Xu*. 2022. Haplotype-resolved de novo assembly of a Tujia genome suggests the necessity for high-quality population-specific genome references. **Cell Systems**. 13 (4), 321-333. e6.

(# co-first author)

Memberships

- Society for Molecular Biology and Evolution