

# Benjamin Chu

954 Fremont St., Apt 1, Menlo Park, CA, 94025, USA

☎ (+1) 909-569-4530 | ✉ [benchu99@hotmail.com](mailto:benchu99@hotmail.com) | 🏠 <https://biona001.github.io/> | 💻 <https://github.com/biona001>

*“Expertise in Mathematics, Statistics, Software Engineering, and Quantitative Biology”*

## Research Interests

<b>Math and Statistics</b>	Knockoffs, Penalized regression methods, Generalized linear models, Variance component models
<b>Computation</b>	Parallel and High-Performance Computing, Big Data Analysis, Package Development
<b>Biology</b>	Polygenic risk scores, Genome-Wide Association Studies, Phasing and imputation, admixture estimation

## Education

### University of California, Los Angeles

Los Angeles, California

PH.D BIOMATHEMATICS

09/2016 - 08/2021

- Advisors: Kenneth Lange and Janet Sinsheimer
- Thesis: Scalable Algorithms for Genetic Association Studies, Genotype Imputation, and Ancestry Inference

### University of California, Berkeley

Berkeley, California

B.A., APPLIED MATHEMATICS

09/2012 - 05/2016

## Positions

### Postdoctoral Researcher, Stanford University

Stanford, California

DEPARTMENT OF BIOMEDICAL DATA SCIENCE (ADVISOR: CHIARA SABATTI)

09/2021 - present

## Experiences

### Google Summer of Code

NumFOCUS - Julia Cohort

STUDENT SOFTWARE DEVELOPER

Summer 2018

- An unique summer program where students work on self-proposed open source projects
- Added 3 additional features to IHT.jl to integrate it with the Open Mendel umbrella program

### Academia Sinica

Institute of Biomedical Sciences

RESEARCH ASSISTANT WITH PROF CARMAY LIM

Summer 2014, 2015, 2016

- One publication: An Efficient Protocol for Computing the pKa of Zn-Bound Water
- Using computational approaches to model drug-protein and drug-environment interactions

## Publications

### Knockoffs.jl: Fast, robust, and easy variable selection for statistical and genetic analysis

work in progress

BB CHU, T MORRISON, C SABATTI

### Fast Score Tests for Multivariate and Longitudinal GWAS with Non-Gaussian Bases

work in progress

BB CHU, JI SS, ZHOU H, SINSHEIMER JS, SABATTI C, LANGE K

### Controlling the False Discovery Rate Improves Polygenic Risk Prediction with sufficient power

work in progress

BB CHU, C SABATTI

## Unsupervised Discovery of Ancestry Informative Markers and Genetic Admixture Proportions in Biobank-Scale Data Sets

S KO, **BB CHU**, D PETERSON, C OKENWA, JC PAPP, DH ALEXANDER, EM SOBEL, H ZHOU, K LANGE

*preprint*

Oct, 2022

## A Flexible Quasi-Copula Distribution for Statistical Modeling

SM JI, **BB CHU**, JS SINSHEIMER, H ZHOU, K LANGE

*preprint*

May, 2022

## Multivariate Genomewide Association Analysis by Iterative Hard Thresholding

**BB CHU**, S KO, JJ ZHOU, A JENSEN, H ZHOU, JS SINSHEIMER, K LANGE

*preprint*

April, 2021

## Optimized Replacement T4 & T3 Dosing in Male & Female Hypothyroid Patients with Different BMIs using a Personalized Mechanistic Model of Thyroid Hormone Regulation Dynamics

M CRUZ-LOYA\*, **BB CHU**\*, J JONKLAAS, DF SCHNEIDER, J DiSTEFANO III (\* DENOTES CO-FIRST AUTHORSHIP)

*Frontiers in Endocrinology*

July, 2022

## A Fast Data-Driven Method for Genotype Imputation, Phasing, and Local Ancestry Inference: MendelImpute.jl

**BB CHU**, EM SOBEL, R WASIOLEK, JS SINSHEIMER, H ZHOU, K LANGE

*Bioinformatics*

July, 2021

## Iterative Hard Thresholding in GWAS: Generalized Linear Models, Prior Weights, and Double Sparsity

**BB CHU**, KL KEYS, CA GERMAN, H ZHOU, JIN J. ZHOU, JS SINSHEIMER, K LANGE

*GigaScience*

June, 2020

## OpenMendel: A Cooperative Programming Project for Statistical Genetics

H ZHOU, JS SINSHEIMER, D BATES, **BB CHU**, CA GERMAN, S JI, KL KEYS, G MOSHER, J PAPP, EM SOBEL, J ZHAI, JJ ZHOU, K LANGE

*Human Genetics*

January, 2020

## An Efficient Protocol for Computing the pKa of Zn-Bound Water

C GRAUFFEL, **BB CHU**, C LIM

*Physical Chemistry Chemical Physics*

November, 2018

## Public Talks

---

### False discovery rate control for polygenic risk prediction

CENTER FOR ADMIXED POPULATIONS AND HEALTH EQUITY (ONLINE)

*University of California, Los Angeles*

April 21, 2022

### Scalable algorithms for genetic association studies, genotype imputation, and ancestry inference

UNIVERSITY OF SOUTHERN CALIFORNIA, BIOSTATISTICS SEMINAR (ONLINE)

*University of Southern California*

January 28, 2021

### Scalable algorithms for GWAS, genotype imputation, and ancestry inference

UNIVERSITY OF MICHIGAN, BIOSTATISTICS SEMINAR (ONLINE)

*University of Michigan*

December 14, 2020

### Julia Meets Mendel: Algorithms and Software for Modern Genomic Data Analysis

AMERICAN SOCIETY OF HUMAN GENETICS VIRTUAL MEETING 2020

*Online*

October 27-30, 2020

### OpenMendel Programming Workshop

INAUGURAL LANGE SYMPOSIUM

*University of California, Los Angeles*

February 22, 2020

### A Multiple Regression Approach for GWAS and High Dimensional Inference

QUANTITATIVE AND COMPUTATIONAL BIOSCIENCES RESEARCH LUNCH

*University of California, Los Angeles*

February 8, 2020

### MendelIHT.jl: Generalized Linear Models for High Dimensional Genetics (GWAS) Data

JULIA CON 2019, AVAILABLE AT: [WWW.YOUTUBE.COM/WATCH?V=UPIKAFShwFW](https://www.youtube.com/watch?v=UPIKAFShwFW)

*University of Maryland Baltimore*

July 25, 2019

## Teachings

---

## Guest Lecturer, Biomath 203

UCLA

### BIOMATH PRECEPTORSHIP

2020

- Taught 2 lectures at graduate probability course (Biomath 203) on random graph theory
- Wrote lecture notes, prepared and graded homework assignments, presented real-time simulation demos
- Materials available at: <https://github.com/biona001/teaching>

## Bruins-In-Genomics (B.I.G) summer research program

UCLA

### GRADUATE STUDENT MENTOR

Summer 2018, 2019, 2020, 2021

- Mentors 2 undergraduates for 8 weeks each summer.
- Proposed student projects. Reviewed student code. Guided them through technical aspects of genetics research.

## Math 98 and 198 course

UC Berkeley

### UNDERGRADUATE STUDENT INSTRUCTOR

2013 - 2016

- Taught 7 semesters of beginner/advanced Rubik's cube course (2 units), with about 15 students per semester.
- Average instructor rating 4.8/5.0

# Honors & Awards

## FELLOWSHIPS AND SCHOLARSHIPS

2019	<b>JuliaCon Financial Assistance for Travel,</b>	<b>Baltimore, MD</b>
2019	<b>Doctoral Student Travel Grant,</b>	<b>Pisa, Italy</b>
2018	<b>JuliaCon Financial Assistance for Travel,</b>	<b>London, UK</b>
2018	<b>Genomics Analysis Training Grant (T32),</b>	<b>NIH/UCLA</b>
2016	<b>University Fellowship,</b>	<b>UCLA</b>
2016	<b>Registration Fee grant,</b>	<b>UCLA</b>

## MISCELLANEOUS

2017	<b>Certified in lifeguarding, first aid, AED, and CPR.,</b>	<b>Amer. Red Cross</b>
2015	<b>International inter-university cube relay, 10th place</b>	<b>Cal Cube Club</b>
2014	<b>UC Berkeley semester Go (Weiqi) Tournament, 1st place</b>	<b>UCB Go Club</b>
2013	<b>3 by 3 rubik's cube speedsolve (10.52 seconds), 144th place in U.S, 910th in world</b>	<b>WCA</b>

# Notable Softwares

<b>Knockoffs.jl</b>	Fast and robust variable selection with Knockoffs
<b>MendelIHT.jl</b>	Efficient $\ell_0$ penalized regression tool for model selection, optimized for gene-mapping (GWAS) analysis
<b>MendelImpute.jl</b>	Fast data-driven genotype imputation, phasing, and admixture estimation
<b>VCFTools.jl</b>	Julia utilities for handling VCF (variant call format) files
<b>Thyrosim.jl</b>	Individualized mechanistic Thyroid hormone simulator for optimizing LT4 and LT4+LT3 dosing
<b>MendelKinship.jl</b>	Computation of theoretical/empirical kinship and other identity coefficients based on pedigree/genotype data
<b>SnpArrays.jl</b>	Package for handling binary PLINK 1 formatted files