

# Zong-Yan Liu

PLANT BREEDING AND GENETICS PHD CANDIDATE

175 Biotechnology Building, 526 Campus Road, Ithaca, NY 14853

[z1843@cornell.edu](mailto:z1843@cornell.edu) | [in zongyanliu](https://www.linkedin.com/in/zongyanliu) | [zongyanliu.github.io/web](https://github.com/zongyanliu)

Ph.D. candidate in Plant Breeding and Genetics at Cornell University (Buckler Lab), with a minor in Computational Biology. I build machine learning and computational genomics tools that make plant genome interpretation more accurate, scalable, and reproducible. My work focuses on end to end pipelines for genome annotation and sequence based modeling, with an emphasis on clean data workflows, robust evaluation, and software that is practical for large cross species analyses.

## Education

### Cornell University

Ithaca, New York

PH.D. IN PLANT BREEDING AND GENETICS

August 2022 - Present

- **Minors:** Computational Biology
- **Advisor:** Edward S. Buckler, Ph.D.

### National Taiwan University

Taipei, Taiwan (R.O.C.)

M.S. IN PLANT BIOLOGY, GPA: 3.963/4.0

September 2018 - June 2020

- **Thesis:** Analysis of *IbHypSys*-mediated MicroRNAs upon Wounding in Sweet Potato (*Ipomoea batatas* cv. Tainung 57)
- **Advisor:** Shih-Tong Jeng, Ph.D., College of Life Science Dean

### National Chung Hsing University

Taichung, Taiwan (R.O.C.)

B.S. IN AGRONOMY, GPA: 3.90/4.0

September 2014 - June 2018

- **Minors:** Management Information Systems and Applied Economics
- **Advisor:** Chin-Ying Yang, Ph.D., Vice President for Student Affairs

## Experience

### Cornell University / USDA-ARS - Buckler Lab

Ithaca, New York

GRADUATE RESEARCH ASSISTANT

August 2022 - Present

- Developed deep learning pipeline to annotate Angiosperm genome for all plant community
- Contributed to the development of a DNA large language foundation model trained on 65 genomes, enabling broad application across all angiosperm plant species
- Utilized a DNA language model to identify and characterize all seed storage proteins across diverse plant species

### Academia Sinica / Institute of Information Science - Bio-IT-Station

Taipei, Taiwan (R.O.C.)

GRADUATE RESEARCH ASSISTANT

August 2021 - July 2022

- Developed and integrated an extrachromosomal circular DNA (eccDNA) database to enable cancer risk score analysis and genomic feature modeling using high-throughput sequencing data.
- Maintained and updated a comprehensive long non-coding RNA (lncRNA) database and website, supporting genomic research and data accessibility for the scientific community.

### National Taiwan University / Machine Learning and Bioinformatics Laboratory - c4Lab

Taipei, Taiwan (R.O.C.)

GRADUATE RESEARCH ASSISTANT

January 2022 - July 2022

- Developed a deep learning-based pipeline to predict human microRNA target genes by integrating sequencing data, transcriptome annotation, biological structure, and microRNA biogenesis features, achieving up to 98% accuracy and outperforming existing tools

### National Taiwan University / Laboratory of Gene Regulation - Jeng Lab

Taipei, Taiwan (R.O.C.)

GRADUATE RESEARCH ASSISTANT

September 2018 - June 2020

- Applied computational and bioinformatics approaches to predict and analyze rice physiological traits regulated by small RNAs and hydrogen peroxide signaling.
- Developed and characterized transgenic sweet potato lines overexpressing primary microRNAs to investigate regulatory networks in stress response and crop improvement.
- Designed, cloned, and introduced target genes using *Agrobacterium*-mediated infiltration in tobacco for functional validation of predicted microRNA targets.
- Predicted and identified 17 novel and 15 conserved microRNAs in sweet potato by integrating small RNA sequencing, transcriptome, and degradome data, revealing their roles in stress and defense responses.

## National Chung Hsing University / Laboratory of Crop Stress Physiology and Signal Delivery – Yang Lab

Taichung, Taiwan (R.O.C.)

UNDERGRADUATE RESEARCH ASSISTANT

February 2017 - June 2018

**Topic:** Ethylene Signaling and Physiological Trait Analysis in Rice (*Oryza sativa*) under Flooding Stress

- Conducted in-depth physiological and molecular analyses of rice under flooding stress, evaluating key traits such as growth, survival, energy conversion, antioxidant activity, and cellular injury
- Collaborated with the Agricultural Experiment Station at the College of Agriculture and Natural Resources, NCHU, to plant and manage four rice strains, and performed systematic sample collection for advanced research studies

## National Cheng Kung University / Institute of Life Science – Huang Lab

Tainan, Taiwan (R.O.C.)

MINISTRY OF SCIENCE AND TECHNOLOGY STUDENT RESEARCH PROGRAM

December 2016 - January 2017

**Topic:** Molecular Studies of the Phytotoxic Potential of *Tolypocladium inflatum* Volatiles on Rice and Weeds)

- Investigated the molecular and physiological effects of *Tolypocladium inflatum*-derived volatile organic compounds (VOCs) on monocot (rice, Poaceae) and dicot (vegetable crop) species, focusing on their potential for selective weed suppression and sustainable agriculture
- Conducted co-culture experiments using VOC-producing GT22 fungal strains with 12 different rice varieties to assess root inhibition, cell apoptosis, reactive oxygen species (ROS) accumulation, and physiological responses
- Applied molecular biology and biochemical assays to identify key genes and regulatory pathways involved in plant responses to fungal VOCs, elucidating the gene regulatory mechanisms behind monocot inhibition
- Collaborated on the development of innovative, nature-based weed management strategies to reduce reliance on chemical herbicides and promote organic farming practices

## Ministry of Agriculture / Agricultural Research and Extension Station

Tainan, Taiwan (R.O.C.)

UNDERGRAD INTERN PROGRAM

June 2016 - August 2016

- Conducted field surveys and phenotypic trait analysis of rice, sesame, soybean, and peanut to support crop improvement and breeding programs

## Achievements & Awards

---

### American Society of Plant Biologists Ambassador

2026

Competitively Selected as a global representative to promote plant biology research, foster community engagement, and bridge communication between the Society and institutions.

### Cornell Exemplary Leadership & Service Award – Early-Career Graduate Students

2024

Campus-wide honor recognizing just 1–3 pre-A-exam scholars each year for exceptional leadership, outreach and inclusion impact across Cornell's 10,000-plus graduate community

### Phi Tau Phi Scholastic Honor Society (R.O.C.), Honorary Member

2020

Invitation-only induction reserved for the top 1 % of bachelor's, top 3 % of master's and top 10 % of doctoral graduates for combined academic and moral excellence

### University "Excellent Performance in Labor Education" Award

2016

Selected as the highest service-learning scores in the university's compulsory experiential-labor curriculum

### Ching-O Award – Outstanding Moral & Intellectual Performance

2015

Single nominee per class approved by college committee for all-round scholarship, leadership and service

### China Youth Corps Exemplary Youth Representative

2013

Nationally selected cohort of emerging leaders invited to represent Taiwanese youth in cultural-exchange and public-service forums

### National Energy Science & Technology Creative Design Contest – Bronze Medal Award

2012

Reached the national finals and secured a prize for an innovative renewable-energy prototype in the Ministry of Education competition

### Certificate

Fundamentals of Deep Learning – Nvidia

2025

## Scholarships & Fellowships

---

**Conference Fellowship, ISMB 2026**, Selected from 1,000+ scientific presentations at ISCB's premier flagship conference.

2026

**Government Studying Abroad Fellowship, Ministry of Education**, National scholarship awarded to 225 recipients out of 1,100+ applicants.

2026

**Government Fellowship for Studying Abroad, Ministry of Education**, a national-level exam with < 1 % success rate

2022 – 2025

**National Key Fields Fellowship for Studying Abroad, Ministry of Education**, allotted just 23 places nationwide

2022 – 2026

**Mu-Lin Scholarship, National Chung Hsing University**, the Department of Agronomy selects just 4 students per semester

2017

**Agricultural Association of Taiwan Scholarship**, the Association caps most awards at ≤ 10 recipients per category annually

2016

**Known-You Social Welfare Foundation Scholarship**, each college may nominate just 2 – 3 students per term

2016

**Prof. Hsin-Sheng Tsay Scholarship**, 2 agronomy majors nationwide are chosen each year

2015

**Tainan City Education Association Cultural & Educational Foundation Scholarship**, limited to 12 awardees each year

2014

**Taiwan Livestock Research Institute Scholarship, Ministry of Agriculture**, high GPA scholarship

2015 - 2020

## Publications

---

### Peer-reviewed Articles

- Ojeda-Rivera JO, Barnes AC, Ainsworth EA, Angelovici R, Basso B, Brindisi LJ, Brooks MD, Busch W, Buttelmann GL, Castellano MJ, Chen J, Costich DE, de Leon N, Emmett BD, Ertl D, Fitzsimmons SL, Flint-Garcia SA, Gore MA, Guan K, Hale CO, Herr S, Hirsch CN, Holding DH, Holland JB, Hsu SK, Hua J, Hufford MB, Kaeppler SM, Leary EN, **Liu Z-Y**, Mahama AA, McCubbin TJ, Messina CD, Michael TP, Miller SJ, Murray SC, Okumoto S, Oren E, Park AN, Piñeros MA, Pugh NA, Raboy V, Rellán-Álvarez R, Romay MC, Rooney T, Roston RL, Sawers RJH, Schnable JC, Schulz AJ, Scott MP, Springer NM, Washburn JD, Zambrano MA, Zhai J, Zou J, & Buckler ES. (2025) **Designing nitrogen-efficient, cold-tolerant maize for modern agricultural systems.** *The Plant Cell* 37 (7): koaf139. <https://doi.org/10.1093/plcell/koaf139>
- Zhai J, Gokaslan A, Schiff Y, Berthel A, **Liu Z-Y**, Lai W, Miller ZR, Scheben A, Stitzer MC, Romay MC, Buckler ES, & Kuleshov V (2025) **Cross-species modeling of plant genomes at single-nucleotide resolution using a pretrained DNA language model.** *Proceedings of the National Academy of Sciences of the United States of America* 122 (24): e2421738122. <https://doi.org/10.1073/pnas.2421738122>
- Lin T-C, Liu Y-L, Liu Y-T, Liu W-H, **Liu Z-Y**, Chang K-L, Chang C-Y, Ni H-C, Huang J-H, & Tsai H-K (2023) **TRIPBASE: a database for identifying the human genomic DNA and lncRNA triplexes.** *NAR Genomics and Bioinformatics* 5 (2): lqad043. <https://doi.org/10.1093/nargab/lqad043>

### Preprints

- Stitzer MC, Seetharam AS, Scheben A, Hsu SK, Schulz AJ, AuBuchon-Elder TM, El-Walid M, Ferebee TH, Hale CO, La T, **Liu Z-Y**, McMorrow SJ, Minx P, Phillips AR, Syring ML, Wrightsman T, Zhai J, Pasquet R, McAllister CA, Malcomber ST, Traiperm P, Layton DJ, Zhong J, Costich DE, Dawe RK, Fengler K, Harris C, Ireland Z, Llaca V, Parakkal P, Zastrow-Hayes G, Woodhouse MR, Cannon EK, Portwood JL II, Andorf CM, Albert PS, Birchler JA, Siepel A, Ross-Ibarra J, Romay MC, Kellogg EA, Buckler ES, & Hufford MB (2025) **Extensive genome evolution distinguishes maize within a stable tribe of grasses.** *bioRxiv* (preprint) 2025.01.22.633974. <https://doi.org/10.1101/2025.01.22.633974>
- Zhai J, Gokaslan A, Hsu SK, Chen S-P, **Liu Z-Y**, Marroquin E, Czech E, Cannon B, Berthel A, Romay MC, Pennell M, Kuleshov V, & Buckler ES (2025) **PlantCAD2: A Long-Context DNA Language Model for Cross-Species Functional Annotation in Angiosperms.** *bioRxiv* (preprint) 2025.08.27.672609. <https://doi.org/10.1101/2025.08.27.672609>
- Liu Z-Y**, Berthel A, Czech E, Stitzer MC, Hsu SK, Pennell M, Buckler ES, & Zhai J (2025) **GeneCAD: Plant Genome Annotation with a DNA Foundation Model.** *bioRxiv* (preprint) 2025.10.31.685877. <https://doi.org/10.1101/2025.10.31.685877>

### Journal Referee Service / Manuscript Reviewing

Plant Physiology and Nature Communication Biology

## Patent

---

- Yang C-Y, Lee C-A, **Liu Z-Y**. Method for Improving the Accuracy of Plant Leaf Temperature Measurement via Thermal Image Analysis. *Republic of China (Taiwan) Invention Patent 1873511*. National Chung Hsing University (assignee). Granted 21 Feb 2025; valid until 15 Jan 2043.

## Presentations

---

### Presentations

- Liu Z-Y**, Berthel A, Czech E, Stitzer MC, Hsu SK, Pennell M, Buckler ES, & Zhai J 2025. **GeneCAD: Plant Genome Annotation with a DNA Foundation Model.** *Cornell Plant Breeding & Genetics Seminar*, Cornell University, Ithaca, NY, USA, 18 November 2025.
- Liu Z-Y**, Berthel A, Gokaslan A, Kuleshov V, & Buckler ES, Zhai J. 2025. **ReelGene2: a large-language-model framework for single-base-pair gene annotation across diverse plant genomes.** Presentation at the *67<sup>th</sup> Annual Maize Genetics Conference*, St Louis, MO, USA, 6–9 March 2025.
- Liu Z-Y**, Zhai J, Stitzer MC, Miller ZR, & Buckler ES. 2024. **Enhancing plant genome annotation with a DNA language model: case studies in Arabidopsis and maize.** Presentation at the *66<sup>th</sup> Annual Maize Genetics Meeting*, Raleigh, NC, USA, 29 February–3 March 2024.
- Liu Z-Y**, & Buckler ES. 2023. **Using machine learning to represent proteins: an example from seed-storage proteins.** *Cornell Plant Breeding & Genetics Seminar*, Cornell University, Ithaca, NY, USA, 17 October 2023.

### Conference posters

- Liu Z-Y**, Berthel A, Czech E, Marroquin E., Stitzer MC, Hsu SK, Pennell M, Buckler ES, & Zhai J 2025. **GeneCAD: Plant Genome Annotation with a DNA Foundation Model.** *The XC Cold Spring Harbor Symposium on Quantitative Biology, AI in Biology*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA, May 26 – Sunday, May 31, 2026.
- Liu Z-Y**, Berthel A, Czech E, Marroquin E., Stitzer MC, Hsu SK, Pennell M, Buckler ES, & Zhai J 2025. **GeneCAD: Plant Genome Annotation with a DNA Foundation Model.** *The 34th International Conference on Intelligent Systems for Molecular Biology*, Washington, D.C., USA, July 12–16, 2026.
- Liu Z-Y**, Berthel A, Czech E, Marroquin E., Stitzer MC, Hsu SK, Pennell M, Buckler ES, & Zhai J 2025. **GeneCAD: Plant Genome Annotation with a DNA Foundation Model.** *Plant Biology 2026 (ASPB/CSPB Joint Conference)*, Ottawa, Canada, July 18–22, 2026.
- Liu Z-Y**, Berthel A, Gokaslan A, Kuleshov V, Buckler ES, & Zhai J. **ReelGene2: a large-language-model framework for single-base-pair gene annotation in diverse plant genomes.** Poster presented at the *67<sup>th</sup> Annual Maize Genetics Conference*; St Louis, MO, USA; 6 – 9 March 2025.
- Liu Z-Y**, Zhai J, Stitzer MC, Miller ZR, & Buckler ES. **Enhancing plant genome annotation with a DNA language model: case studies in Arabidopsis and maize.** Poster presented at the *66<sup>th</sup> Annual Maize Genetics Meeting*; Raleigh, NC, USA; 29 February – 3 March 2024.
- Liu Z-Y**, Berthel A, Zhai J, & Buckler ES. **Transforming plant gene annotation through a DNA foundation model.** Poster presented at the *34th International Conference on Arabidopsis Research*; San Diego, CA, USA; 15 – 19 July 2024.
- Liu Z-Y**, Zhai J, & Buckler ES. **Identification of seed-storage proteins across *Andropogoneae* genomes with machine learning.** Poster presented at the *65<sup>th</sup> Annual Maize Genetics Meeting*; St Louis, MO, USA; 16 – 19 March 2023.
- Liu Z-Y**, & Jeng S-T. **Identification and functional characterization of *l*bHypSys-mediated microRNAs upon wounding in sweet potato.** Poster presented at *Plant Biology 2020 (American Society of Plant Biologists)*; Washington, DC, USA; 25 – 29 July 2020. doi: 10.46678/PB.20.1007170

## Teaching and Service

---

### Teaching Experience

#### MENTOR, NSF REU Program – Boyce Thompson Institute (BTI)

Summer 2026

Supervise undergraduate researchers in deploying GeneCAD for standardized plant genome annotation and benchmarking evolutionary signatures across species using Python-based comparative genomics workflows.

#### MENTOR, AI Genomics Fellowship – U.S. Department of Agriculture, Agricultural Research Service (ARS)

2025 – present

Provide one-on-one mentorship to AI fellows, offering specialized guidance on algorithm development, data analysis pipelines

#### CONVERSATION FACILITATOR, Language Resource Center – Cornell University

2024 – present

Design weekly speaking drills, cultural modules, and formative assessments for beginning and intermediate learners

#### INSTRUCTOR, “Biomedical Data Mining” – National Taiwan University

Summer 2021

Delivered lectures on conda environment management, RNA-seq pipelines, BLAST searches, Galaxy workflows, and building sequence databases.

#### INSTRUCTOR, “Methods in Plant Molecular Biology Research” – National Taiwan University

Spring 2020

Taught a module on gene-feature discovery and functional annotation in plants; developed problem sets and graded practical reports.

### Professional & Academic Service

#### PROGRAM COMMITTEE MEMBER - Machine Learning in Computational and Systems Biology, Intelligent Systems for Molecular Biology

2026

Selected to serve at ISMB, contributing to the peer review and evaluation of research submissions at the intersection of machine learning and computational biology.

#### ADVISORY BOARD MEMBER - Cornell International Services Advisory Board, Cornell University

2025 – 2026

Advise on policy and programming for international scholars; coordinate feedback between academic units and International Services.

#### SYSTEM MAINTENANCE ENGINEER - Institute of Plant Biology, National Taiwan University

2018 – 2022

Designed and maintained the institute’s official website

Administered Google Business profiles for the institute and NTU cafeterias, ensuring up-to-date public information

#### EXECUTIVE OFFICER - NCHU project “The Effects of Climate Change and the Analysis of Possible Adaptation Strategies”

2017 – 2018

Assisted Asst. Prof. Yi-Yuan Su in course logistics and student support

Drafted official correspondence with Taiwan’s Ministry of Education to invite international scholars and coordinate project reporting

## University Leadership & Governance

---

### Cornell Graduate and Professional Student Assembly (GPSA)

Ithaca, New York

Cornell’s GPSA serves as the official governance body representing 10k+ graduate and professional students to the university administration.

BIOLOGICAL SCIENCES REPRESENTATIVE (2024 – PRESENT) · FINANCE COMMISSION SECRETARY (2025 – PRESENT) · GRADUATE SCHOOL DEAN-SEARCH COMMITTEE MEMBER (2025) · SECRETARY, DIVERSITY & INTERNATIONAL STUDENTS COMMITTEE (2024 – PRESENT)

- Act as liaison between the GPSA, Cornell Graduate School, and university administration, circulating policy updates and gathering feedback from graduate and professional students across the sciences
- As Finance Commission Secretary, review and approve yearly GPSA budgets and grant requests—including support for Cornell Cinema and Anabel’s Grocery—ensuring transparent allocation of student activity funds
- Organized campus-wide cultural programs and information sessions that ease the transition of international students and address their concerns
- Represented the graduate student body on the Cornell Graduate School Dean-Search Committee, helping evaluate candidates and formulate selection recommendations

## Student Organizations

---

### Synapsis (Cornell Plant Breeding Graduate Student Association)

Ithaca, New York

MEMBER, CO-PRESIDENT (2024 – 2025) · INVITED SPEAKER CO-LEADER (2023 – 2024) · SYMPOSIUM COMMITTEE (2022 – 2023)

2022 – 2025

- Direct overall programming for the association, coordinating cultural events, NAPB-2024 outreach, and professional-development workshops.
- Co-led the annual graduate-student recruitment weekend, managing logistics and peer-mentoring activities.
- Served on the 2022 Symposium committee, delivering a seamless hybrid (in-person + virtual) experience for presenters and attendees.
- Organized seminar visits—handling speaker invitations, schedules, and on-campus hosting—to expand departmental networking and learning opportunities.

### Cornell Taiwanese Student Association (CTSA)

Ithaca, New York

PRESIDENT (2023 – 2024)

- Acted as the official liaison linking Taiwan’s Ministry of Foreign Affairs, Cornell Taiwanese alumni, and graduate students.
- Guided new Taiwanese students through visa applications, housing, banking, phone service, and mental-health resources.
- Coordinated campus-wide cultural and networking events in partnership with the undergraduate chapter and the Ithaca Taiwanese Association.
- Developed the CTSA website and a comprehensive student handbook in collaboration with Cornell International Services and the Office of Global Learning.