JASON WILLIAMS

Office: 516-367-5187 Email: williams@cshl.edu

Education

2004	**Stony Brook University, State University of New York—Stony Brook, New York** Bachelors of Science, Biology

Professional Appointments

2025-Present	**Assistant Director** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2021-2025	**Assistant Director, Inclusion and Research Readiness** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2016-2021	**Assistant Director, External Collaborations** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2013-2024	**Lead, CyVerse (formerly iPlant) Education, Outreach, and Training** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2011-2012	**Manager, iPlant Collaborative Education, Outreach, and Training** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2008-Present	**Lead Instructor** Science Institute Yeshiva University High School for Girls, Hollis, NY
2008-2011	**Educator** DNA Learning Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2007-2008	**Laboratory Technician** Trotman Lab, Cancer Center Cold Spring Harbor Laboratory, Cold Spring Harbor NY
2004-2007	**Laboratory Technician** Lukowitz Lab, Plant Science Cold Spring Harbor Laboratory, Cold Spring Harbor NY

Professional Development

2014 Comparative Genomics Course—Cold Spring Harbor Laboratory
2007 Immunocytochemistry, In Situ Hybridization & Live Cell Imaging—Cold Spring Harbor Laboratory

Professional Society Memberships

American Association for the Advancement of Science
International Society for Computational Biology
American Society for Microbiology
American Society of Plant Biologists

Professional Service and Advisory Roles

2024 - Present	Member, NIH Council of Councils GREI Working Group, National Institutes of Health
2023 - Present	Member, Education Advisory Committee, Allen Institute
2023 - Present	Member, Scientific Advisory Board, HudsonAlpha Institute for Biotechnology
2023 - Present	Member, External Advisory Board, NSF ACCESS-CI Computing
2023 - Present	Member, Learning and Training, Wellcome Connecting Science, Wellcome Trust
2023 - Present	Member, EDI Committee, International Society for Computational Biology
2022 - Present	Chair (2022-2024), Secretary (2020-2022), Board of Directors, Dryad Data Repository
2022 - 2024	Member, Data Works Advisory Committee, Federation of Sci. Societies for Experimental Biology
2022 - Present	Member, Scientific Advisory Board, Lupus Research Alliance
2022 - 2023	Chair, Kavli Frontiers of Sci. 2023, US Symposium, National Academy of Sciences
2021 - 2022	Member, Org. Comm., Kavli Frontiers of Sci. 2022, US Symposium, National Academy of Sciences
2021 - Present	Member, DEIJ Committee, Earth BioGenome Project
2021 - Present	Member, Organizing Committee, Bioinformatics Open Source Conference
2021 - Present	Editor, Journal of Open Source Education
2020 - 2023	Chair, Training Advisory Group, Australian BioCommons
2019 - 2020	Curriculum Advisory Board, Health Data Research UK
2018 - Present	Member, External Expert Panel, BioData Catalyst, National Institutes of Health-NHLBI
2018 - Present	Member, Science and Industry Advisory Board, ELIXIR UK
2017 - 2019	Member, Steering Committee, Plante, American Society of Plant Biology
2017 - 2019	External Panel of Consultants, Data Commons Pilot Project, OD, National Institutes of Health
2016 - 2019	Chair, International Science Advisory Group: EMBL Australian Bioinformatics Resource
2016 - 2017	Chair, Steering Committee, Software Carpentry Foundation
2015 - 2017	Treasurer, Steering Committee, Software Carpentry Foundation
2013 - 2015	Member, Scientific Training Advisory Board, The Genome Analysis Centre (Now Earlham Institute)

Review Service

Foundations and Organizations

Biotechnology and Biological Sciences Research Council—UK Research and Innovation
National Aeronautics and Space Administration
National Open Research Forum, Ireland
NIH National Cancer Institute

NIH National Human Genome Research Institute
U.S. National Science Foundation

Conferences and Journals

Bioinformatics
Bioinformatics Open Source Conference
Pacific Symposium on Biocomputing
Plant Cell
Plant Direct
PLOS Biology
PLOS Computational Biology

Awards and Honors

2025	Elizabeth W. Jones Award for Excellence in Education, Genetics Society of America
2024	Genome—Unlocking Life's Code Exhibit—Black History in Genetics: Trailblazers, Pioneers, and Leaders in Genetics; NIH NHGRI/Smithsonian Museum
2020 - 2023	Kavli Frontiers of Science Fellow, National Academies of Science
2020	Meritorious Winner—NSF 2026 Idea Machine, National Science Foundation

Presentations and Teaching

Note: Workshops and invited workshops, unless otherwise noted, are events I organized and/or served as instructor for.

2025

- **Invited Speaker: Nanopore Network — Broadening Participation in the Benefits of Genomics** *International Plant and Animal Genome Meeting (PAG 33), San Diego, CA*

2024

- **Session Chair/Organizer: Advanced Computational Methods—CyVerse for Machine Learning, Containers, and Clouds** *International Plant and Animal Genome Meeting (PAG 31), San Diego, CA*
- **Session Organizer/Speaker: Earth BioGenome Project—The Missing Components in Genomics: Justice, Equity, Diversity, Inclusion** *International Plant and Animal Genome Meeting (PAG 31), San Diego, CA*
- **Invited Speaker: Decoding DNA** *Cocktails and Chromosomes Outreach Event, Huntington, NY*
- **Invited Speaker: Effective and Inclusive Professional Development** *Effectively Communicating Bioimage Analysis, East Sussex, UK*
- **Invited Speaker: Effective and Inclusive Professional Development** *Center for Scientific Collaboration and Community Engagement, Online*
- **Workshop: Nanopore for Educators Workshop** *NIH Outreach Event, J.F.K. High School, Tamuning, Guam*
- **Invited Organizer/Speaker: Inclusive Genomics Education—Empowering Everyone to Sequence Everything** *American Society of Human Genetics, Online*
- **Summit Chair/Session Organizer: Bicycle Principles** *6th Global Bioinformatics Education Summit, New York, NY*
- **Workshop: Nanopore for Educators** *Spelman College: Summer of Nanopore Sequencing Workshop, Atlanta, GA*

- **Workshop: Nanopore for Educators** *Arecibo C3 STEM Center: Summer of Nanopore Sequencing Workshop, Arecibo, PR*
- **Workshop: Nanopore for Educators** *U. Hawaii, Manoa: Summer of Nanopore Sequencing Workshop, Honolulu, HI*
- **Workshop: 2024 NHGRI Short Course in Genomics—Nanopore Hands-On for Educators** *NIH National Human Genomics Research Institute, Bethesda, MD*
- **Workshop: Computational Basics for Plant Biology** *Frontiers and Techniques in Plant Science, Cold Spring Harbor, NY*
- **Invited Participant: The Next Steps for Vision and Change** *American Association for the Advancement of Science, Washington, DC*
- **Session Chair/Organizer: Holistic Genomics Research Experiences—A Path to Equity** *Fulfilling the Promise of Nanopore Sequencing in Education, Huntington, NY*
- **Instructor** *CSHL Course - Advanced Sequencing Technologies & Bioinformatics Analysis, Cold Spring Harbor, NY*
- **Invited Talk: It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences** *Westminster College, London, UK*

2023

- **Session Chair/Organizer: Advanced Computational Methods—CyVerse for Machine Learning, Containers, and Clouds** *International Plant and Animal Genome Meeting (PAG 30), San Diego, CA*
- **Session Organizer: A Moderated Discussion on the Missing Components in Genomics—Justice, Equity, Diversity, Inclusion** *International Plant and Animal Genome Meeting (PAG 30), San Diego, CA*
- **Invited Speaker: Understanding Barriers to Bioinformatics Education - Network for Integrating Bioinformatics into Life Sciences Education** *International Plant and Animal Genome Meeting (PAG 30), San Diego, CA*
- **Invited Keynote: Professionalizing Training—Origin Stories for the Modern Researcher** *Australian Research Data Commons Skills Summit, Online*
- **Instructor** *CyVerse Foundational Open Science Skills, Online*
- **Invited Panelist: Researcher Perspectives—NIH Generalist Repository Ecosystem Initiative (GREI) Workshop** *NIH Office of Data Science, Generalist Repository Ecosystem Initiative (GREI) Workshop, Online*
- **Invited Panelist: Career-Spanning Learning in the Life Sciences - Supporting Open, Data-Driven Biology** *OpenScapes—Pathways to Open Science, Online*
- **Invited Speaker: Training and Re-Training in Open and Reproducible Research** *University of Florida, Rigor and Reproducibility Seminar Series, Online*
- **Invited Fellow/Organizing Committee Chair; Organizer: Proteins, Biology, and Artificial Intelligence** *2023 U.S. Kavli Frontiers of Science Symposium, Irvine, CA*
- **Invited Participant** *CZI—Accelerating Open Science in Latin America, Buenos Aires, AR*
- **Invited Keynote: It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences** *AgBioData Annual Conference, Chicago, IL*
- **Invited Seminar: It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences** *HudsonAlpha, Huntsville, AL*
- **Invited Speaker: Building Communities of Practice** *Bioinformatics Education Summit, Hinxton, UK*
- **Organizer/Instructor** *Nanopore Sequencing Foundations for Course-Based Research, New York, NY*
- **Invited Keynote: Making Life Science CUREs Inclusive and Accessible** *Gordon Research Conference: Undergraduate Biology Education Research, Lewiston, ME*

- **Organizing Committee** *Bioinformatics Open Source Conference, Lyon, FR*
- **Instructor** *CSHL Course—Advanced Sequencing Technologies & Bioinformatics Analysis*
- **Invited Seminar: It's Impossible to Keep Up—Career-Spanning Learning in the Life Sciences** *North Carolina State University, Genetics and Genomics Academy, Raleigh, NC*

2022

- **Workshop Discussant: Workshop on Support for Open Science Training and Education** *Chan Zuckerberg Initiative, Online*
- **Workshop: Plotting and Programming in Python** *National Society of Black Engineers, Anaheim, CA*
- **Invited Keynote: Career-Spanning Learning in the Life Sciences—Supporting Open, Data-Driven Biology** *North-East Regional Meeting Society for Developmental Biology, Woods Hole, MA*
- **Invited Speaker: Learning New Tricks—Data Sharing and Career-Spanning Learning in STEM** *Federation of American Societies for Experimental Biology DataWorks!, Online*
- **Invited Fellow/Organizing Committee; Organizer: COVID-19 Vaccine Development and Therapeutic RNA Molecules** *2022 U.S. Kavli Frontiers of Science Symposium, Irvine, CA*
- **Invited Lecture: Open Science and Learning in the Genome Age** *National Academies of Science, Irvine, CA*
- **Organizer/PI: Making Career-Spanning Learning in the Life Sciences Inclusive and Effective for All** *CSHL Banbury Center, Cold Spring Harbor, NY*
- **Invited Attendee, Session Moderator: NINDS Workshop on Catalyzing Communities of Research Rigor Champions** *National Institutes of Health, Bethesda, MD*
- **Instructor:** *CyVerse Foundational Open Science Skills, Tucson, AZ*
- **Invited Workshop: Computational Basics for Plant Biology** *Frontiers and Techniques in Plant Science, Cold Spring Harbor, NY*
- **Conference Organizer/Panel Host/BOSC and Education COSI Joint-Keynote: "Riding the Bicycle — Including All Scientists on a Path to Excellence"; Panel Organizer - Building and Maintaining Inclusive Open Science** *Bioinformatics Open Source Conference, Madison, WI*
- **Workshop: 2021 NHGRI Short Course in Genomics—DNA Barcoding** *NIH National Human Genomics Research Institute, Online*
- **Invited Speaker: Opening the Door to Good Science Education** *Johns Hopkins Welsh Symposium on R3 Reforming Graduate Science Education, Online*
- **Invited Speaker: Open Science—The Next Generation** *NASA Ames Horizons in Biosciences and Informatics Seminar Series, Online*
- **Invited Speaker: Career-Spanning Learning in the Life Sciences —Supporting Open, Data-Driven Biology** *Southwest Society for Developmental Biology Meeting, Austin, TX*
- **Invited Workshop: Computational Basics for Plant Biology** *Advanced Sequencing Technologies & Bioinformatics Analysis, Cold Spring Harbor, NY*
- **Invited Keynote: Running to Stand Still - Traversing Skill Gaps in Computational Biology** *Australian Bioinformatics and Computational Biology Society Annual Meeting, Melbourne, AU*

2021

- **Instructor:** *CyVerse Foundational Open Science Skills, Online*
- **Invited Speaker: Northeast Student Data Corps (NSDC) Data Science Career Panel** *NSF NorthEast Big Data Hub, Online*
- **Workshop: Introduction to Bioinformatics with CyVerse** *North Carolina State University, Online*
- **Workshop: Genomics Data Carpentry** *Atlanta University Center Consortium Data Science Initiative, Online*

- **Invited Workshop: Introduction to RNA-Seq with the Kallisto and Sleuth Workflows** *SfAM International Applied Microbiology Conference 2021, Online*
- **Workshop: JMU Center for Genome and Metagenome Studies (CGEMS)—Getting Started with R and CyVerse** *James Madison University, Online*
- **CyVerse Tools for Reproducible Research** *Plant Biology 2021, Online*
- **Invited Speaker: Extending the Possibilities of Biology with Open Science** *Prairie View A&M University College of Engineering, Online*
- **Invited Keynote: 2021 NHGRI Short Course in Genomics - Owning the Genome** *NIH National Human Genomics Research Institute, Online*
- **Invited speaker: Relevance and Rapport—Promoting Inclusion Through Public Engagement** *American Society of Cell Biology Public Engagement Webinar, Online*
- **Opening Doors—Preparing the Next Generation of Computational Biologists** *Black Women in Computational Biology, Online*
- **Webinar: CyVerse Visual Interactive Computing Environment** *CyVerse Webinar Series, Online*
- **Workshop: Integrating Bioinformatics into Your Courses** *American Society of Cell Biology, Cell Bio Virtual 2021, Online*

2020

- **Session Chair/Organizer: Advanced Computational Methods—Machine Learning, Containers, and Clouds** *International Plant and Animal Genome Meeting (PAG XXVIII), San Diego, CA*
- **Session Chair/Organizer: CyVerse —Software, Tools, and Services for Data-Driven Discovery, Data Science, and Education** *International Plant and Animal Genome Meeting (PAG XXVIII), San Diego, CA*
- **Invited Speaker: Resources and Programs for Undergraduate Education in Genomics - "Tools and Approaches for Making Bioinformatics Work in the Classroom"** *International Plant and Animal Genome Meeting (PAG XXVIII), San Diego, CA*
- **Session Organizer: Pragmatic Solutions for Scaling Your Analysis - Machine Learning, Imaging, Containers, Clouds, and APIs** *International Plant and Animal Genome Meeting (PAG XXVIII), San Diego, CA*
- **Invited Speaker: Cloud Computing for Genomics Data Analysis —"Reproducibility in the Clouds —Harnessing the *in nube* Paradigm"** *International Plant and Animal Genome Meeting (PAG XXVIII), San Diego, CA*
- **Awardee Talk: NSF 2026 Idea Machine Award Ceremony —Reinventing Scientific Talent** *National Science Foundation, Alexandria, VA*
- **Instructor: Introduction to Reproducible Science** *CyVerse Foundational Open Science Skills (FOSS) Camp, Tucson, AZ*
- **Selected Talk: The Genomics Education Alliance —Scalable, Sustainable Infrastructure for Undergraduate CUREs** *The Allied Genetics Conference 2020, Online*
- **Selected Talk: Reproducible Analyses with Jupyter and R on CyVerse** *Plant Biology 2020 Online, Online*
- **Selected Workshop: RNA-Seq and Jupyter - Bringing Genomics Into the Classroom** *2021 Biology and Mathematics Educators (BIOME) Institute, Online*
- **Poster: Making Bioinformatics Tools Classroom-Friendly** *2021 Biology and Mathematics Educators (BIOME) Institute, Online*
- **Invited Fellow/Participant** *2020 U.S. Kavli Frontiers of Science Symposium, Online*
- **Invited Seminar: Computational Biology** *Carnegie Mellon University, Online*
- **Invited Ecology Graduate Seminar** *University of Arizona, Online*

2019

- **Session Chair/Organizer: Advanced Computational Methods—Machine Learning, Containers, and Clouds** *International Plant and Animal Genome Meeting (PAG XXVII), San Diego, CA*
- **Session Chair/Organizer: CyVerse - Software, Tools, and Services for Data-Driven Discovery, Data Science, and Education** *International Plant and Animal Genome Meeting (PAG XXVII), San Diego, CA*
- **Session Chair: Resources and Programs for Undergraduate Education in Genomics** *International Plant and Animal Genome Meeting (PAG XXVII), San Diego, CA*
- **Session Organizer: Pragmatic Solutions for Scaling Your Analysis: Machine Learning, Imaging, Containers, Clouds, and APIs** *International Plant and Animal Genome Meeting (PAG XXVII), San Diego, CA*
- **Invited Workshop: Extensible Cyberinfrastructure for Life Science** *5th Annual Meeting of the Arkansas Bioinformatics Consortium, Little Rock, AR*
- **Lightning Talk: Broadening Participation with Bioinformatics, Big Data, and Data Science** *Environmental Data Science Inclusion Network, Boulder, CO*
- **Workshop Developer** *CyVerse Foundational Open Source Skills Camp, Tucson, AZ*
- **Workshop Organizer/Instructor** *Genomics Data Carpentry - Genomics 2.0 Lesson Release Workshop, Tucson, AZ*
- **Frontiers in Plant Science Course, Invited Instructor: Introduction to R** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Invited Opening Keynote: Education COSI - "Training, Teaching, Technology, Togetherness – Promoting Knowledge Exchange in Life Sciences Through Communities of Practice"** *International Society for Computational Biology: Joint Annual Meeting - 27th Intelligent Systems for Molecular Biology/18th European Conference on Computational Biology, Basel, CH*
- **Invited Speaker: Scalable Plant-Research in Cloud Environments - "Overview of CyVerse Tools and Services: Introduction to Data/Metadata Management and Sharing with CyVerse"** *International Society for Computational Biology: Joint Annual Meeting - 27th Intelligent Systems for Molecular Biology/18th European Conference on Computational Biology, Basel, CH*
- **Lightning Talk: The Data Carpentry Genomics Curriculum: Overview and Impact** *International Society for Computational Biology: Joint Annual Meeting - 27th Intelligent Systems for Molecular Biology/18th European Conference on Computational Biology, Basel, CH*
- **Invited Seminar: "Cyberinfrastructure for Scaling Research, Education, and People"** *Swiss Institute of Bioinformatics, Lausanne, CH*
- **Invited Talk: LifeSciTrainers.org: A Global Community of Practice for Short-Format Training in the Life Sciences** *GOBLET Annual Meeting, Jakarta, ID (Virtual)*
- **CyVerse Austria - Train the Trainers** *Technical University of Graz, Graz, Austria*
- **Invited Seminar: Computational Learning, Thinking, and Doing in 21st Century Biology** *University of Puerto Rico, San Juan, Puerto Rico*
- **CyVerse Tools, Services, and Reproducibility Workshop** *University of Puerto Rico, San Juan, Puerto Rico*
- **Invited Keynote: The View from Nowhere in Computational Infrastructure** *Elixir UK All-Hands, University of Dundee, Dundee, Scotland*

2018

- **Poster: Broadening Participation with Bioinformatics, Big Data, and Data Science** *NSF INCLUDES Summit, Alexandria, VA*
- **Session Chair/Organizer: CyVerse Education: DNA Subway, Data Science, and Microbial Genomics - CyVerse - Software, Tools, and Services for Data-Driven Discovery; BOF-Organizer: Machine Learning,**

Imaging, Containers, Clouds and APIs *International Plant and Animal Genome Meeting (PAG XXVI), San Diego, CA*

- **Genomics Data Carpentry and CyVerse Tools and Services Workshops** *North Carolina State University, Raleigh, NC*
- **Symposium Keynote Speaker: "Data, Data Everywhere, Nor Any a Drop to Drink"** *ASPB Southern Section Regional Meeting, New Orleans, LA*
- **Workshop: Data Carpentry: Ecology** *Chan-Zuckerberg Biohub, San Francisco, CA*
- **Workshop: Genomics Data Carpentry and CyVerse Tools and Services Workshops** *University of New Orleans, New Orleans, LA*
- **Invited Seminar: "Scaling Science by Scaling People, Purpose-built Cyberinfrastructure for the Life Sciences"** *Amazon Research, Cambridge, UK*
- **Workshop: Genomics Data Carpentry and Carpentry Con** *University College, Dublin, IE*
- **Joint Session Talk: "Improving the Undergraduate Bioinformatics Curriculum", Training and Documentation in Bioinformatics** *Bioinformatics Open Source Conference/Galaxy Community Conference, Portland, OR*
- **Instructor: Frontiers in Plant Science Course, Invited Instructor: Genomics Data Carpentry - Introduction to R** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Workshop: Genomics Data Carpentry and Machine Learning** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Poster: "The Next 10 Years of Data-Driven Plant Biology"/ Data Carpentry and RNA-Seq Workshop** *Plant Biology 2018, Montreal, CN*
- **Genomics Data Carpentry and CyVerse Tools and Services Workshops** *Colorado State University, Ft. Collins, CO*
- **Poster: "Improving the Bioinformatics Curriculum"** *CSHL Biological Data Science, Cold Spring Harbor, NY*
- **Invited Keynote: "Improving the Bioinformatics Curriculum", Invited Instructor: "Introduction to Kallisto and Sleuth RNA-Seq Workflows - Bulk and Single-cell RNA-Seq Applications"** *BioInfoSummer, Perth, AU*
- **Workshop: CyVerse UK RNA-Seq workshop** *Earlham Institute, Norwich, UK*

2017

- **Session Chair/Organizer: CyVerse Education: DNA Subway, Data Science, and Microbial Genomics - CyVerse - Software, Tools, and Services for Data-Driven Discovery** *International Plant and Animal Genome Meeting (PAG XXV), San Diego, CA*
- **Workshop: CyVerse UK Tooks and Services Workshop** *University of York, York, UK*
- **Symposium Chair: Computation/Bioinformatics: CyVerse Resources for Enabling Data-driven Plant Biology; Panelist: Common Author Errors and How to Avoid Them; ASPB Journal Editor's Session: Tools for Reproducible Plant Biology** *Plant Biology 2017 - American Society of Plant Biology, Honolulu, HI*
- **Instructor: Genomics Data Carpentry** *Cold Spring Harbor Laboratory - Frontiers & Techniques in Plant Science, Cold Spring Harbor, NY*
- **Panelist: Careers in Education and Government** *Beyond the Bench 2017, Cold Spring Harbor Laboratory*
- **Workshop: Genomics Data Carpentry and CyVerse Tools and Services Workshops** *University of Arkansas, Fayetteville, AK*
- **Workshop: Genomics in Education and CyVerse Tools and Services Workshops** *James Madison University, Harrisonburg, VA*
- **Workshop: Software Carpentry** *Los Alamos National Laboratory, Los Alamos, NM*

- **Workshop: Genomics Data Carpentry and CyVerse Tools and Services Workshops** *New Mexico State University, Las Cruces, NM*
- **Panelist: Plantae Conviron Scholars Career Panel** *ASPB Conviron Scholars, Online*
- **Invited Speaker: TorBUG Bioinformatics Seminar —"The Bioinformatics Training Landscape"** *Toronto Bioinformatics User Group, Toronto, CN*

2016

- **Session Chair/Organizer: iPlant and Education: DNA Subway, Data Science, and Microbial Genomics - Cyberinfrastructure for Life Science and Beyond – Scaling your Science with iPlant** *International Plant and Animal Genome Meeting (PAG XXIV), San Diego, CA*
- **Workshop: Genomics Data Carpentry** *Stony Brook University, Genomics Cluster, Stony Brook, NY*
- **Workshop: Genomics in Education Workshop/CyVerse Tools and Services Workshop** *University of Delaware, Newark, DE*
- **Invited Workshop Organizer: CyVerse Tools for Data Driven Discovery** *MidSouth Computational Biology and Bioinformatics (MCBIOS), Memphis, TN*
- **Workshop: CyVerse Tools and Services Workshop** *Ohio State University, Wooster, OH*
- **Webinar: Education and Outreach Webinar Series: iPlant's DNA Subway** *iDigBio, Online*
- **Invited Lecture/Workshop Organizer: CyVerse Training Day** *University of Melbourne, Melbourne, AU*
- **Workshop: CyVerse and Atmosphere Cloud Overview** *National eResearch Collaboration Tools and Resources Project (Nectar), Melbourne, AU*
- **Instructor: Big Genomic Data Skills Training for Professors: Invited Lecture/Workshop Organizer: CyVerse Cloud for Research and Education** *Jackson Laboratory, Farmington, CT*
- **Workshop: Genomics Data Carpentry/BD2K** *National Institutes of Health, Bethesda, MD*
- **Workshop: Genomics Data Carpentry/CyVerse Tools and Services** *Langebio/Cinvistav, Irapuato, MX*
- **Guest Lecture/Workshop Organizer: CyVerse Overview/CyVerse Atmosphere for Education** *2016 National Academies Special Topics Summer Institute on Quantitative Biology/BioQUEST, Raleigh, NC*
- **Invited Talk: Bioinformatics Symposium: Resources for Big Data Biology: The Evolution of iPlant** *Plant Biology 2016, Austin, TX*
- **Workshop Organizer: CyVerse Tools and Services** *Genetics Society of America/Allied Genetics Conference, Orlando, FL*
- **Participant: NSF RCN Kick-off Meeting** *Network for Integrating Bioinformatics into Life Science Education (NIBLSE), Omaha, NE*
- **Keynote Speaker: Leveraging Cyberinfrastructure to Scale Science and People** *South African Genetics Society/South African Bioinformatics Society - Joint Conference, Durban, ZA*
- **Workshop: Genomics Data Carpentry Workshop** *NorthWest University, Potchefstroom, ZA*
- **Invited Instructor: Accelerating Plant Science with CyVerse** *CSHL Cereal Genomics Workshop, Cold Spring Harbor, NY*
- **Poster: Required Parameters - What Does it Take to Bring Bioinformatics into the Classroom at a National Level?** *Biological Data Science, Cold Spring Harbor, NY*
- **Invited Lecture/Workshop Organizer: EMBL-ABR Workshop on Open Source and Software Development Best Practice** *EMBL-Australian Bioinformatics Resource, Melbourne, AU*
- **Invited Lecture/Workshop Organizer: EMBL-ABR Open and Scalable Training Workshop** *EMBL-Australian Bioinformatics Resource, Melbourne, AU*

2015

- **Session Chair/Organizer: DNA Subway: Genomics, DNA Barcoding, and RNA-Seq Made Easy for the Undergraduate Classroom - iPlant Workshop: Why a Biology Cyberinfrastructure is the Best Way to Meet the Challenges of Large Datasets** *International Plant and Animal Genome Meeting (PAG XXIII), San Diego, CA*
- **Workshop: iPlant Tools and Services Workshop** *Clemson University, Clemson, SC*
- **Workshop: Data Carpentry Hackathon on Genomics Lesson Materials** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Workshop: USDA-iPlant BigData Workshop on Genome Variation** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Workshop: iPlant Tools and Services Workshop** *University of Georgia, Athens, GA*
- **Workshop: iPlant Tools and Services and Genomics in Education Workshop** *Santa Fe Community College, Santa Fe, NM*
- **Workshop: RNA-Seq for the Next Generation Workshop** *Bowie State University, Bowie, MD*
- **Workshop: iPlant Tools and Services and Genomics in Education Workshop** *South Dakota State University, Brookings, SD*
- **Workshop: RNA-Seq for the Next Generation Workshop** *California State University, Long Beach, Long Beach, CA*
- **Workshop: iPlant Tools and Services Workshop** *The Genome Analysis Centre (Earlham Institute), Norwich, UK*
- **Invited Panelist: Open Source, Open Door: Increasing Diversity in the Bioinformatics Open Source Community** *Bioinformatics Open Source Conference, Dublin, IE*
- **Workshop: Using Biological Cyberinfrastructure to Scale Science and People – Applications in Data Storage, HPC, Cloud Analysis, and Bioinformatics Training** *ISMB/European Conference on Computational Biology, Dublin, IE*
- **Invited Talk: Bioinformatics Symposium: How to Scale Data, Science, and People Using Biological Cyberinfrastructure - PUI Workshop: Using iPlant Cyberinfrastructure to Support Diverse Research and Education Objectives** *Plant Biology 2015, Minneapolis, MN*
- **Workshop: Data Carpentry Workshop at Plant Biology 2015** *Plant Biology 2015, Minneapolis, MN*
- **Invited Takl: Overview of iPlant Resources Cyberinfrastructure for Metagenomics** *USDA-ARS Meeting on Microbial Genomics, Biomics, and Metagenomics Workshop, Athens, GA*
- **Workshop: iPlant Genomics in Education Workshop** *Florida SouthWestern College, Ft. Myers, FL*
- **Workshop: iPlant Multiworkshop with Data-Intensive-Biology Lab** *University of California, Davis, Davis, CA*
- **Invited Speaker: MIP Seminar: How to Scale Data, Science, and People Using Biological Cyberinfrastructure** *Colorado State University, Ft. Collins, CO*
- **Invited participant: NEON/Data Carpentry Hackathon on Geospatial Data** *National Ecological Observatory Network, Boulder, CO*
- **Workshop: iPlant Workshop at CSHL Plant Genomes and Biotechnology Meeting** *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- **Invited Talk: Scaling Plant Science (and Scientists) Using Biological Cyberinfrastructure** *Plant Genomes & Biotechnology: From Genes to Networks, Cold Spring Harbor, NY*

2014

- **Session Chair/Organizer: DNA Subway Workshop - iPlant Workshop: Cyberinfrastructure for All Life Sciences**
  International Plant and Animal Genome Meeting (PAG XXII), San Diego, CA
- **Workshop: iPlant Tools and Services Workshop**
  Linfield College, McMinnville, OR
- **Invited Seminar: The iPlant Collaborative - A Unified Cyberinfrastructure for the New Life Science**
  The Genome Analysis Centre (Currently: Earlham Institute), Norwich, UK
- **Workshop: The iPlant Collaborative and CoGe**
  Wageningen University, Wageningen, NL
- **Workshop: Introduction to iPlant Collaborative**
  2014 Maize Meeting, Beijing, CN
- **Workshop: iPlant Train-the-Trainers/Tools and Services Workshop**
  UC Davis, Davis, CA
- **Invited Participant: NSF (RCN) Incubator: Network for Integrating Bioinformatics into Life Sciences Education (NIBLSE)**
  University of Omaha, Omaha, NE
- **Workshop: Software Carpentry Workshop**
  Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- **Workshop: iPlant Tools and Services Workshop**
  Stony Brook University, Stony Brook, NY
- **Workshop: Understanding GWAS Workshop**
  University of Nebraska, Lincoln, NE
- **Workshop: Symposium Speaker: CyVerse Resources for Plant Science**
  Plant Biology 2014, Portland, OR
- **Workshop: Software Carpentry Workshop**
  Plant Biology 2014, Portland, OR
- **Workshop: iPlant Tools and Services/Genomics in Education Half-Day Workshops**
  Botany 2014, Boise, ID
- **Invited Talk: DNA Subway – An Educational Bioinformatics Platform for Gene and Genome Analysis: DNA Barcoding, and RNA-Seq**
  World Congress on Genetics Applied to Livestock Production, Vancouver, CA
- **Invited Talk: NPGI Annual Awardee Meeting – The iPlant Collaborative: Computational Resources for Young Investigators - How Can We Help?**
  National Science Foundation, Arlington, VA
- **Invited Participant: High Performance Computing (HPC) in Undergraduate Biology Education**
  Cold Spring Harbor Laboratory (Banbury), Cold Spring Harbor Laboratory, NY
- **Invited Workshop: Dual Benefit of Bioinformatics Training**
  European Conference on Computational Biology, Strasbourg, FR
- **Workshop: ARS Arthropod Genomics Workshop**
  USDA-ARS, Beltsville, MD
- **Meeting Founder/Organizer/Session Chair: Biological Data Science (Inaugural Meeting)**
  Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- **Workshop: iPlant Train-the-Trainers**
  Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Other teaching, mentoring, and programs

- **Science field trips and summer courses, DNA Learning Center (2008-Present)**

    - High school summer courses created
        - **BioCoding (2013-Present)**
            - This camp introduces the basic skills necessary for bioinformatics and will equip motivated students with knowledge that will serve them well beyond course participation. Enrolled students should bring an introductory knowledge of biology, but little to no knowledge of computer programming or bioinformatics is required. I created this course to bring bioinformatics to our summer high school programs.
        - **Sequence-a-genome (2020-Present)**
            - This course uses small, affordable Nanopore DNA sequencers—which plug into a computer through a USB drive—to do real-time DNA sequencing. Students will generate complete, composite sequences of their own mitochondrial genomes and the class will "adopt" an organism to sequence as a research project. They will make use of bioinformatics and data science to assemble the genome sequences and explore what DNA reveals about humans and other living things. The class will also explore genomics applications, ethics, and the revolutionary effects of high throughput sequencing in health, agriculture, and other fields. I created this course to bring hands-on genomics to our summer high school programs.
    - High school programs created
        - **Science and Technology Research Scholars (2019-Present)**
            - STARS is a two-week summer research experience designed to promote full participation in the next generation of scientists, doctors, and other health professionals. This program provides students with state-of-the-art laboratory and computer science skills needed to succeed in STEM (science, technology, engineering, mathematics) in college and beyond, and tailored to the needs of students from contexts that may have excluded them from STEM. I created this program in collaboration with Dr. Carol Carter of Stony Brook University.
- **Science Institute at Yeshiva University for Girls (2008-Present)**
    - I created and have been the instructor for this 1.5-3 hour per-week course over the fall and spring semesters at Yeshiva University High School. This course is designed with the intention of producing thinkers first and scientists second. The fall portion of the course will give a broad introduction of the skills in modern molecular biology. Students will cover basic experiments and techniques in molecular biology. In the spring semester, students put skills into scientific context with additional application of their work to semi-independent research projects. I created this course upon request from the high

school. It is inspired by my high school AP biology course, and teacher, Mr. Raymond McGraime, who was influential. It was and continues to be my way of giving back.

- **Professional programs created**

    - **Foundational Open Science Skills (2020-2023 instructed; course ongoing)**
        - Foundational Open Science Skills is a multi-week online course built as part of CyVerse, and designed primarily for graduate students and early-career faculty. The goal was to familiarize learners with practices and approaches for data-intensive research. The curriculum featured progressive engagement through hands-on exercises and interactive discussions on topics such as FAIR data management, reproducible science, and open science practices. Activities included developing and reviewing data management plans, authoring reproducible workflows in Python, working with software containerization, and managing data-intensive projects.
    - **LifeSciTrainers.org (2019-Present)**
        - LifeSciTrainers.org is a ~700 online community of practice for anyone who delivers short-format training (workshops, boot camps, short-courses, etc.) – often on new and quickly evolving topics – in the life sciences. Community-of-practice involves establishing forums for these trainers to share their knowledge, resources, and expertise, develop and promote standards for training, open educational resources, and establish the role of trainer as a vital position within the life science community. Monthly calls and discussions are a way to highlight and strengthen this underappreciated role within the life science community. I founded this community and recruited a steering committee to advance these goals.
    - **Genomics Data Carpentry (2015-Present)**
        - Genomics Data Carpentry is a curriculum designed for workshops that equip learners with essential skills for working with genomics data, including data management and software tools for genomics research. The curriculum emphasizes best practices for organizing bioinformatics projects and data, using command-line utilities, analyzing sequence quality, performing variant calling, and connecting to and utilizing cloud computing resources.I hosted the 2015 kickoff for the group of instructional co-developers and led the finalization and initial rollout of the lessons later that year. The curriculum continues to be actively used and updated, and I currently maintain the R-based genomics lessons.
    - **CSHL Biological Data Science Meeting (2014-Present)**
        - The scope of the Biological Data Science meeting will be the infrastructure, software, and algorithms needed to analyze large data sets in biological research. We welcome abstracts from researchers in both academia and industry who work on technical aspects of the discussion topics in all areas of

biology, from genomics to imaging. We also welcome abstracts from translational and clinical researchers who regularly mine large data sets as part of their projects. The goal is to assemble a multidisciplinary audience that will discuss best practices, identify challenges, and highlight successes in the analysis of large biological data sets. This meeting was started based upon a proposal I submitted to CSHL Meetings and Courses.

- **Professor-for-a-day (2008-Present)**
  - The purpose of Cold Spring Harbor Laboratory's "Professor For a Day" program is to show students what science involves outside the laboratory, in particular how scientists communicate their work and subject it to peer review. In a structured group setting, students will attend talks and a poster session at one of the many scientific meetings hosted each year by Cold Spring Harbor Laboratory. While participating in top science fairs may give students a taste of what a career in science is like, participating in "Professor For a Day" will offer a genuine sense of how the scientific community operates. This mentorship program was started based upon a proposal I submitted to CSHL Meetings and Courses.

Publications and Products

Peer reviewed publications

1. Reed LK, Kleinschmit AJ, Buonaccorsi V, Hunt AG, Chalker D, **Williams JJ**, Jones CJ, Martinez-Cruzado JC, Rosenwald A. A genomics learning framework for undergraduates. PLoS One. 2025;20(1):e0313124. doi: 10.1371/journal.pone.0313124. eCollection 2025. PubMed PMID: 39787200; PubMed Central PMCID: PMC11717232.
2. Drew J, Morgan W, Galindo S, Kleinschmit A, McWilliams M, Pauley M, Triplett E, **Williams JJ**, Murdoch B, Rosenwald A. Revisiting barriers to implementation of bioinformatics into life sciences education. Front Educ. 2023 Nov;8:-. doi: 10.3389/feduc.2023.1317191.
3. **Williams JJ**, Tractenberg RE, Batut B, Becker EA, Brown AM, Burke ML, Busby B, Cooch NK, Dillman AA, Donovan SS, Doyle MA, van Gelder CWG, Hall CR, Hertweck KL, Jordan KL, Jungck JR, Latour AR, Lindvall JM, Lloret-Llinares M, McDowell GS, Morris R, Mourad T, Nisselle A, Ordóñez P, Paladin L, Palagi PM, Sukhai MA, Teal TK, Woodley L. An international consensus on effective, inclusive, and career-spanning short-format training in the life sciences and beyond. PLoS One. 2023;18(11):e0293879. doi: 10.1371/journal.pone.0293879. eCollection 2023. PubMed PMID: 37943810; PubMed Central PMCID: PMC10635508.
4. Swetnam TL, Antin PB, Bartelme R, Bucksch A, Camhy D, Chism G, Choi I, Cooksey AM, Cosi M, Cowen C, Culshaw-Maurer M, Davey R, Davey S, Devisetty U, Edgin T, Edmonds A, Fedorov D, Frady J, Fonner J, Gillan JK, Hossain I, Joyce B, Lang K, Lee T, Littin S, McEwen I, Merchant N, Micklos D, Nelson A, Ramsey A, Roberts S, Sarando P, Skidmore E, Song J, Sprinkle MM, Srinivasan S, Stanzione D, Strootman JD, Stryeck S, Tuteja R, Vaughn M, Wali M, Wall M, Walls R, Wang L, Wickizer T, **Williams JJ**, Wregglesworth J, Lyons E.

CyVerse: Cyberinfrastructure for open science. *PLoS Comput Biol.* 2024;20(2):e1011270. doi: 10.1371/journal.pcbi.1011270.

5. Handelsman J, Elgin S, Estrada M, Hays S, Johnson T, Miller S, Mingo V, Shaffer C, **Williams JJ**. Achieving STEM diversity: Fix the classrooms. Science. 2022 Jun 3;376(6597):1057-1059. doi: 10.1126/science.abn9515. Epub 2022 Jun 2. PubMed PMID: 35653460.
6. Harris NL, Cock PJA, Fields CJ, Hokamp K, Maia J, Munoz-Torres M, Sharan M, **Williams JJ**. BOSC 2021, the 22nd Annual Bioinformatics Open Source Conference. *F1000Res.* 2021;10. doi: 10.12688/f1000research.74074.1. eCollection 2021. PubMed PMID: 34721839; PubMed Central PMCID: PMC8524300.
7. llwood ER, Sessa JA, Abraham JK, Budden AE, Douglas N, Guralnick R, Krimmel E, Langen T, Linton D, Phillips M, Soltis PS, Studer M, White LD, **Williams JJ**, Monfils AK. Biodiversity Science and the Twenty-First Century Workforce. *Bioscience.* 2020 Feb 1;70(2):119-121. doi: 10.1093/biosci/biz147. Epub 2019 Dec 18. PubMed PMID: 32063647; PubMed Central PMCID: PMC7012771.
8. Nüst D, Eddelbuettel D, Bennett D, Cannoodt R, Clark D, Daróczi G, Edmondson M, Fay C, Hughes E, Kjeldgaard L, Lopp S, Marwick B, Nolis H, Nolis J, Ooi H, Ram K, Ross N, Shepherd L, Sólymos P, Swetnam TL, Lee T, Turaga N, Van Petegem C, **Williams JJ**, Willis C, Xiao N. The Rockerverse: Packages and Applications for Containerisation with R. *R J.* 2020;12(1):437-. doi: 10.32614/RJ-2020-007.
9. Clarke DJB, Wang L, Jones A, Wojciechowicz ML, Torre D, Jagodnik KM, Jenkins SL, McQuilton P, Flamholz Z, Silverstein MC, Schilder BM, Robasky K, Castillo C, Idaszak R, Ahalt SC, **Williams JJ**, Schurer S, Cooper DJ, de Miranda Azevedo R, Klenk JA, Haendel MA, Nedzel J, Avillach P, Shimoyama ME, Harris RM, Gamble M, Poten R, Charbonneau AL, Larkin J, Brown CT, Bonazzi VR, Dumontier MJ, Sansone SA, Ma'ayan A. FAIRshake: Toolkit to evaluate the FAIRness of research digital resources. *Cell Syst.* 2019 Nov 27;9(5):417-421. doi: 10.1016/j.cels.2019.09.011. Epub 2019 Oct 30. PubMed PMID: 31677972; PubMed Central PMCID: PMC7316196.
10. **Williams JJ**, Drew JC, Galindo-Gonzalez S, Robic S, Dinsdale E, Morgan WR, Triplett EW, Burnette JM 3rd, Donovan SS, Fowlks ER, Goodman AL, Grandgenett NF, Goller CC, Hauser C, Jungck JR, Newman JD, Pearson WR, Ryder EF, Sierk M, Smith TM, Tosado-Acevedo R, Tapprich W, Tobin TC, Toro-Martínez A, Welch LR, Wilson MA, Ebenbach D, McWilliams M, Rosenwald AG, Pauley MA. Barriers to integration of bioinformatics into undergraduate life sciences education: A national study of US life sciences faculty uncover significant barriers to integrating bioinformatics into undergraduate instruction. *PLoS One.* 2019;14(11):e0224288. doi: 10.1371/journal.pone.0224288. eCollection 2019. PubMed PMID: 31738797; PubMed Central PMCID: PMC6860448.
11. Wilson Sayres MA, Hauser C, Sierk M, Robic S, Rosenwald AG, Smith TM, Triplett EW, **Williams JJ**, Dinsdale E, Morgan WR, Burnette JM 3rd, Donovan SS, Drew JC, Elgin SCR, Fowlks ER, Galindo-Gonzalez S, Goodman AL, Grandgenett NF, Goller CC, Jungck JR, Newman JD, Pearson W, Ryder EF, Tosado-Acevedo R, Tapprich W, Tobin TC, Toro-Martínez A, Welch LR, Wright R, Barone L, Ebenbach D, McWilliams M, Olney KC, Pauley MA. Bioinformatics core competencies for undergraduate life sciences education. *PLoS One.* 2018;13(6):e0196878. doi: 10.1371/journal.pone.0196878. eCollection 2018. PubMed PMID: 29870542; PubMed Central PMCID: PMC5988330.
12. Harper L, Campbell J, Cannon EKS, Jung S, Poelchau M, Walls R, Andorf C, Arnaud E, Berardini TZ, Birkett C, Cannon S, Carson J, Condon B, Cooper L, Dunn N, Elsik CG, Farmer A, Ficklin SP, Grant D, Grau E, Herndon N, Hu ZL, Humann J, Jaiswal P, Jonquet C, Laporte MA, Larmande P, Lazo G, McCarthy F, Menda N, Mungall CJ, Munoz-Torres MC, Naithani S, Nelson R, Nesdill D, Park C, Reecy J, Reiser L, Sanderson LA, Sen TZ, Staton M, Subramaniam S, Tello-Ruiz MK, Unda V, Unni D, Wang L, Ware D, Wegrzyn J, **Williams JJ**, Woodhouse M, Yu J, Main D. AgBioData consortium recommendations for sustainable genomics and

genetics databases for agriculture. *Database (Oxford).* 2018 Jan 1;2018. doi: 10.1093/database/bay088. PubMed PMID: 30239679; PubMed Central PMCID: PMC6146126.

13. Barone L, **Williams JJ**, Micklos D. Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. *PLoS Comput Biol.* 2017 Oct;13(10):e1005755. doi: 10.1371/journal.pcbi.1005755. eCollection 2017 Oct. PubMed PMID: 29049281; PubMed Central PMCID: PMC5654259.
14. **Williams JJ**, Teal TK. A vision for collaborative training infrastructure for bioinformatics. *Ann N Y Acad Sci.* 2017 Jan;1387(1):54-60. doi: 10.1111/nyas.13207. Epub 2016 Sep 7. Review. PubMed PMID: 27603332.
15. Martin L, Cook C, Matasci N, **Williams JJ**, Bastow R. Data mining with iPlant: a meeting report from the 2013 GARNet workshop, Data mining with iPlant. *J Exp Bot.* 2015 Jan;66(1):1-6. doi: 10.1093/jxb/eru402. Epub 2014 Oct 17. PubMed PMID: 25326627.
16. Hilgert U, McKay S, Khalfan M, **Williams JJ**, Ghiban C, Micklos D. DNA Subway. *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment.* XSEDE '14: Annual Conference of the Extreme Science and Engineering Discovery Environment; 13–18 Jul 2014; Atlanta, GA, USA. New York, NY, USA: ACM; 2014.
17. **Williams JJ**, McKay S, Khalfan M, Hilgert U, Lauter S, Jeong ES. DNA Subway—An educational bioinformatics platform for gene and genome analysis: DNA barcoding, and RNA-Seq. *Proceedings of the 10th World Congress on Genetics Applied to Livestock.*

Book chapters

1. **Williams JJ**. Driving Without a License: Career Advice from a Non-PhD Scientist. In: Markovac J, Barrett KE, Garrison H, editors. Life Science Careers [Internet]. Cham, Switzerland: Springer; American Physiological Society; 2024. Available from: https://doi.org/10.1007/978-3-031-50694-9_22.
2. **Williams J**, Nash B, Ghiban C, Khalfan M, Hilgert U, Lauter S, Yang CH, Micklos DA. Analysis of DNA Barcodes Using DNA Subway. Methods Mol Biol. 2024;2744:551-560. doi: 10.1007/978-1-0716-3581-0_34. PMID: 38683342.
3. **Williams JJ**. CyVerse for Reproducible Research: RNA-Seq Analysis. Methods Mol Biol. 2022;2443:57-79. doi: 10.1007/978-1-0716-2067-0_3. PubMed PMID: 35037200.

Preprints (not otherwise published)

1. Elgin S, Hays S, Mingo V, Shaffer C, **Williams JJ**. Building Back More Equitable STEM Education: Teach Science by Engaging Students in Doing Science. *[preprint].* 2021 Jun. doi: 10.1101/2021.06.01.446616.

Other products

1. Website/Bioinformatics Platform: https://edna.dnabarcoding101.org/ , Carol Henger and Jason Williams, DNA Learning Center, Cold Spring Harbor, NY. 2024.
2. Survey Instrument: **Williams JJ**, Buonaccorsi V, Tractenberg R, Chalker DL, Wiley E, Hunt AG, Rosenwald A, Reed L, Subramanya S, Leung W. GEA Bioinformatics Tools Databases Faculty Professional Development survey and GEA Bioinformatics Infrastructure Support survey. 2022 Dec. Available from: https://figshare.com/articles/presentation/GEA_Bioinformatics_Tools_Databases_Faculty_Professional_Development_survey_and_GEA_Bioinformatics_Infrastructure_Support_survey/21789905.
3. Curriculum: Becker EA, Teal T, Michonneau F, Sane R, Reiter T, **Williams JJ**, Charbonneau A, Kweskin M, Zohaib Anwar M, Cadzow M, Martinez PA. Data Carpentry: Genomics Workshop Overview. *Zenodo*; 2019 Jun. Available from: https://datacarpentry.github.io/genomics-workshop/.

4. Website/Bioinformatics Platform: https://dnasubway.cyverse.org/ (Peer reviewed publication 16,17); 2010.

In-preparation

1. Nanopore for Educators eBook; https://nanopore4edu.org/latest/ (2025)
2. DNA sequence as a blueprint for STEM education and the bioeconomy (2025)

Grant Funding Support

Note: Projects listed if I contributed to grant writing and/or served in PI or Co-PI role

2024-2028 **Pathogen Data Network**

The mission of the Pathogen Data Network is to enable a world-wide ecosystem of linked data and tools to support research and public health response to infectious diseases and epidemics. The DNALC will contribute to the onboarding of educators and researchers onto an online bioinformatics platform for the exploration and deposition of pathogen-related data. The DNALC will contribute to national and international workshops and outreach efforts.
Funder: National Institutes of Health, NIAID, RFA-AI-23-032 (Swiss Institute of Bioinformatics)
PIs: Jason Williams (PI, CSHL).
Award: $954,304 (DNALC Component)

2024 **Conference: Global Bioinformatics Education Summit 2024 ? Energizing Communities to Power the Bioeconomy Workforce**

The Global Bioinformatics Education Summit (GBES) 2024, held at the New York Genome Center, addressed critical challenges in bioinformatics education, aiming to build a data-literate life sciences workforce and strengthen the bioeconomy. The Summit brought together 270 participants (70 in person and 200 virtual) from over 50 countries, fostering a global network of educators, industry leaders, funders, and policymakers. A major highlight was the inaugural Funders, Industry, and Policymakers Forum, which promoted the development of a U.S. national bioinformatics education strategy as a cornerstone for driving innovation and competitiveness in the bioeconomy.
Funder: National Science Foundation, DBI 2421267
PIs: Jason Williams (PI).
Award: $47,078

2023-2028 **Collaborative Research: ACSER: Arecibo C3 - Center for Culturally Relevant and Inclusive Science Education, Computational Skills, and Community Engagement**

This award establishes a science, technology, engineering, and mathematics (STEM) education and research center at the federally owned site in Arecibo, Puerto Rico, home to the Arecibo Observatory. The Arecibo Center for Science Education, Computational Skills, and Community Engagement (Arecibo C3) will integrate science (ciencia), computation (computación), and community (comunidad)—geographical, disciplinary, and cultural —to promote relevant and accessible teaching, science education research, outreach, and workforce development. Arecibo C3 is a collaboration of Cold Spring Harbor Laboratory, the University of Puerto Rico at Río Piedras, and the University of Maryland, Baltimore County. The Center is purpose-built to expand access to STEM opportunities and enable broad participation in STEM activities. STEM education

researchers, students, professional educators, and the public will benefit from the Center's education research programs, skill-building initiatives, and workforce development efforts.
Funder: National Science Foundation, DRL 2321759
PIs: Jason Williams (PI), Patricia Ordóñez, Jose Agosto-Rivera.
Award: $7,578,158 (DNALC Component: $2,038,173)

2022-2024 **Enhancing DNA Subway 2.0 as a Shared Resource for Bioscience Workforce Development**
Data science is one of the ten fastest growing careers in the US, and the outlook is also bright for bioinformatics technicians. However, there is a significant gap in the bioinformatics skills of students entering emerging, multidisciplinary, bioscience industries. With initial funding from the NSF CyVerse cyber-infrastructure for biology research, DNA Subway was developed in 2010 as a learner-centered set of online tools to introduce students to bioinformatics and big data analysis. This project will update DNA Subway 2.0 as a more flexible, accessible, and capable resource to prepare students for the modern bioscience workforce.
Funder: National Science Foundation, DUE 2301207
PIs: David Micklos, **Jason Williams (Co-PI)**.
Award: $703,269

2022-2024 **Developing Foundations for Nanopore DNA Sequencing Course-based Undergraduate Research Experiences at Minority-Serving Institutions**
This project aims to serve the national interest by making it easier for educators to integrate advanced DNA sequencing technologies and related laboratory and data science skills into their teaching. In this project, the Cold Spring Harbor Laboratory DNA Learning Center will work with three partner institutions to reduce these barriers to using advanced DNA sequencing technology in the classroom. Together with New York City College of Technology, Spelman College, and University of Puerto Rico, Río Piedras, the project team will co-develop protocols and teaching tools that integrate advanced DNA sequencing into hands-on training for biology undergraduates. Faculty from these institutions and other partners include 2- and 4-year, Hispanic-serving, and Historically Black institutions, maximizing the potential that the solutions developed can be successfully implemented in a variety of college settings. Independent evaluation of the project will lay a foundation for future dissemination at a national level and will support efforts to create a highly skilled and diverse US STEM workforce.
Funder: National Science Foundation, DUE 2216349
PIs: Jason Williams (PI), Jose Agosto-Rivera, James Melton, Jeremy Seto.
Award: $299,348 (DNALC Component: $179,574)

2022-2027 **Genomics STEP-UP High School Coordinating Center Region A; Increasing Diversity Through Genomics Research Experiences**
This project will create a coordinating center to deliver mentoring and research experiences in STEM for high school students from groups historically excluded in the sciences (e.g., ethnic and gender minorities, rural students, economically disadvantaged) across Department of Health & Human Services Region A. Mentors from across the region will be trained and equipped by the Cold Spring Harbor Laboratory DNA Learning Center to bring advanced -omics into a summer research and school year mentoring program. Ultimately, a sustainable network of mentors and students will create new, inclusive opportunities that diversify the nation's STEM workforce.
Funder: National Institutes of Health, NIDDK, 5R25DK132965

PIs: Jason Williams (PI)
Award: $1,173,753

2021-2022 **NSF 2026: What Works in Workshops – Evolving Short-format Training to Serve Life Sciences STEM Professionals in the 21st Century**
This conference explores challenges in applying established principles from education research and professional development, identifies research opportunities, and defines strengths and weaknesses in short-format training. The conference organizers (representing the largest life science short-format training programs world-wide) developed an outline of discussion topics, and a process for identifying participants and incorporating broad community input beyond the initial set of attendees. The output of the meeting includes relevant research questions, sets of lessons learned, and recommendations for supporting career-long learning that will be applicable in the life sciences and other STEM fields. Ultimately, these outputs form a blueprint for improvement that supports communities, organizations, and programs developing and delivering short-format training.
Funder: National Science Foundation, DRL 2027025
PIs: Jason Williams (PI), Rochelle Tractenberg.
Award: $99,774 (DNALC Component: $76,984)

2018-2022 **RCN-UBE: Establishing a Genomics Education Alliance: Steps toward Sustainability**
The GEA will form an extended network, recruiting members from existing undergraduate genomics education communities, and exploit their combined experience to create a unified and sustainable platform, available through a web portal, that faculty can use to engage diverse students in inquiry-based curricula and CUREs. Our efforts will focus around three major aims: 1) curation and maintenance of computational resources for genomics instruction; 2) development of faculty training and student learning materials; and 3) evaluation of teaching and learning assessments in genomics education, aligning assessments with faculty training and student learning resources.
Funder: National Science Foundation, DBI 1827130 (Washington University in St. Louis)
PIs: Douglas Chalker, Anne Rosenwald, Vincent Buonaccorsi, Rochelle Tractenberg, **Jason Williams (Co-PI)**.
Award: $494,749 (DNALC Component: $39,971)

2018-2024 **iPlant Collaborative: A Cyberinfrastructure for Plant Sciences**
The DNALC led CyVerse community engagement efforts including workshops, webinars, and collaborative user support. DNALC will also provide continuing technical support, maintenance, and updating for DNA Subway and other CyVerse educational resources.
Funder: National Science Foundation, DBI 1743442, (University of Arizona)
PIs: David Micklos, John Fonner, Tyson Swetnam, Parker Antin, Eric Lyons, Matthew Vaughn, Nirav Merchant, Eric Lyons.
Award: $15,199,324 (DNALC Component: $805,000)

2013-2017 **Infrastructure and Training to Bring Next-generation Sequence (NGS) Analysis into Undergraduate Education**
This three-year project is developing a sustainable infrastructure and training program to assist undergraduate faculty in integrating next-generation sequence analysis into course-based and

independent student research. Participating faculty develop their own RNA sequence datasets that bear on novel research problems in eukaryotic genomics.
Funder: National Science Foundation, DUE 1323522
PIs: David Micklos.
Award: $538,000

2008-2018 **iPlant Collaborative: A Cyberinfrastructure for Plant Sciences**
As lead of Education, Outreach, and Training (EOT) for this project, the DNALC developed *DNA Subway*, a bioinformatics workflow for education, and training teachers and researchers to use the *iPlant* cyberinfrastructure.
Funder: National Science Foundation, DBI 1265383, (University of Arizona)
PIs: Doreen War, Matthew Vaugh, Nirav Merchant, Stephen Goff, Daniel Stanzione, Eric Lyons.
Award: $50,300,000 (DNALC Component: $6,631,033)