John Didion | PhD

📞 (919) 536 9924 • ✉ johndidion@gmail.com • 🌐 john.didion.net

in jdidion • 🐦 jdidion • ⬡ jdidion • ORCiD 0000-0002-8111-6261

Computational Genomics Innovator

Results-Oriented Research • **Sequencing Expertise** • **Applied Machine Learning** • **Professional Software Engineering**

Creative and highly productive computational biologist with a passion for leveraging high-throughput computing and advanced machine and deep learning to interpret and exploit complex genomic and clinical datasets.

- Next-generation sequencing expert with broad experience in assay design, data processing, bioinformatics, and computational analytics development.
- Talented engineer with 15+ years experience developing production-grade software using Python, R, SQL, C++, Ruby, and Java.
- Effective and award-winning communicator who has won competitive research grants and authored >20 peer-reviewed publications.
- Versatile team member who excels as a leader, collaborator, or individual contributor.

Experience

DNAnexus   Rockville, MD (remote) | 2018 to present
Principal Scientist - FDA Specialist
- Helping to guide development of the PrecisionFDA platform.

Personal Genome Diagnostics   Baltimore, MD | 2017 to 2018
Principal Bioinformatics Scientist, Research and Development
- Lead effort to develop a bioinformatics pipeline to support offering the company's IVD assays on a new sequencing platform, thereby substantially expanding market opportunities.
- Developed a novel software method for NGS error correction that enables accurate low-level variant detection in liquid biopsy samples.

National Human Genome Research Institute   Bethesda, MD | 2014 to 2017
Postdoctoral Fellow, Laboratory of Dr. Francis S Collins
- Leveraged a large, multi-omics dataset to investigate epigenetic mechanisms underlying regulatory variants implicated in type 2 diabetes.
- Developed machine-learning approaches for imputation of missing data in multi-omics experiments.
- Investigated cell-to-cell variability in pancreatic islets, including responses to environmental perturbation, using single-cell gene expression and chromatin accessibility data.
- Designed a novel sequencing assay for single-molecule resolution transcriptome analysis.
- Developed Atropos, user-friendly software for QC and pre-processing of NGS reads. Created reproducible benchmark pipeline using software containers for accompanying publication.
- Initiated collaborative project to undestand genomic diversity of biofilm communities.
- Awarded six grants, including an American Diabetes Association fellowship and an NIH K22.

American Academy of Bioinformatics   Bethesda, MD | 2016 to Present
Instructor of Bioinformatics
- Developed comprehensive, open-source course materials for workshops in DNA-Seq and RNA-Seq data analysis.
- Taught workshops and earned highly positive student reviews.

University of North Carolina at Chapel Hill   Chapel Hill, NC | 2009 to 2014
Research Assistant, Laboratory of Dr. Fernando Pardo-Manuel de Villena
- Characterized a novel meiotic drive locus, *R2d2*, and multiple modifier loci responsible for extreme transmission distortion in interspecific crosses.
- Conducted a GWAS of wild mice to identify genes associated with the accumulation of Robertsonian translocations.
- Developed CLASP, a software tool for validation of cell lines used in research.

Institute for Systems Biology   Seattle, WA | 2007 to 2008
Computational Biology Software Engineer, Laboratory of Dr. Ruedi Aebersold
- Created TIQAM, a work flow management system to support Multiple Reaction Monitoring (MRM) proteomics experiments.

Muze, Inc.   Seattle, WA | 2004 to 2007
Software Development Engineer
- Implemented key components of web services platform for purchase and distribution of digital media, including consumer management and security.
- Created an intelligent installation system that decreased deployment time for the web services platform from days to less than an hour.
- Improved team efficiency by implementing a code generation framework that produced a large percentage of the domain and persistence code for the web services platform.

Encyclopaedia Britannica, Online Services   Chicago, IL | 2003 to 2004
Software Developer
- Developed and tested releases of several new web products, including Spanish- and Chinese- language editions of the company's flagship product.
- Elimiated substantial software licensing costs by migrating marketing and e-commerce systems from outsourced to internal solutions.

ThoughtWorks, LLC   Chicago, IL | 2001 to 2003
Software Developer/Consultant
- Designed, developed, and enhanced user interface, business logic, and persistence-layer components of large-scale financial software package for the commercial leasing industry.
- Trained members of India development team and helped transition project to an international, round-the-clock effort.
- Co-lead efforts to build web services interoperability lab. Developed Java tools for automated compatibility testing of web service runtime environments.

Technical Skills

Programming Languages	*Expert:* Python, R
	Proficient: C++, Java, SQL, Ruby
Machine Learning	*Python:* scikit-learn, keras, networkX
	R: xgboost, e1071, caret
Data Science/Visualization	*Python:* Numpy, Pandas, Seaborn
	R: ggplot2, other "tidyverse" packages
High-Performance Computing	*Containerization:* Docker, Singularity
	Pipelines: Nextflow, Snakemake, CWL
	Job Scheduling: SGE, SLURM, LSF
Genomics	*NGS Assays:* DNA-Seq, RNA-Seq, Methyl-Seq, ATAC-Seq, ChIP-Seq, HiC, Single-Cell (10X, Fluidigm)
	Bioinformatics: Samtools, BWA, GATK
	Other: SNP and methylation array analysis

Education

Doctor of Philosophy, Bioinformatics and Computational Biology	2008 to 2014
University of North Carolina, Chapel Hill, NC	
Bachelor of Science, Computer Science	1996 to 2001
Northwestern University, Evanston, IL	

Honors & Awards

NIH 1 K22 ES028024-01 BD2K Career Transition Award	2017 to Present
A Big Data Approach to Learning the Type 2 Diabetes Regulome	
Career transition award with 3 years tenure-track funding	
American Diabetes Association Postdoctoral Fellowship	2017 to Present
A Multi-Tissue and Multi-Omics Investigaton of Type 2 Diabetes	
Postdoctoral fellowship with up to 3 years salary and research support	
NIH Intramural Sequencing Center Pilot Grants	2014 to 2016
Four separate project proposals funded	
Institutional award with funding for sequencing services	
Department of Health and Human Services Ignite	2014
LabGenius: The Smart Lab Notebook for Scientists	
3-month incubator program to fund innovative projects within HHS	
Dean's Distinguished Dissertation Award (Department Nominee)	2014
University of North Carolina at Chapel Hill	
Nominee from the Bioinformatics and Computational Biology program	
Verne Chapman Young Scientist Award	2013
International Mammalian Genome Society	
Best trainee talk at the International Mammalian Genome Conference	
Chicago Prize	2013
Complex Traits Consortium	
Best graduate student talk at the Complex Traits Consortium meeting	
Genome Research Award for Outstanding Poster	2010
International Mammalian Genome Society	
Outstanding poster at the International Mammalian Genome Conference	

Leadership & Service

Certified Software Carpentry instructor	2016 to Present
Co-taught mutliple workshops on biological data science to novices.	
Organizer, NHGRI Preprint Journal Club	2016 to Present
Started journal club to review and provide feedback on scientific preprints.	
Hackathon team leader	2016 to 2017
Lead teams in prototyping novel bioinformatics tools in multiple hackathons organized by NCBI.	
Graduate and undergraduate student mentor	2013 to Present
Designed and oversaw student projects that lead to peer-reviewed publications.	
Secretariat member (honorary), International Mammalian Genome Society	2013 to 2015