

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

Personal Genome Diagnostics

Baltimore, MD | 2017 to Present

Principal Bioinformatics Scientist, Research and Development

- Leading 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 understand 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.
- Eliminated 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	<i>Expert:</i> Python, R <i>Proficient:</i> C++, Java, SQL, Ruby
Machine Learning	<i>Python:</i> scikit-learn, keras, networkX <i>R:</i> xgboost, e1071, caret
Data Science/Visualization	<i>Python:</i> Numpy, Pandas, Seaborn <i>R:</i> ggplot2, other "tidyverse" packages
High-Performance Computing	<i>Containerization:</i> Docker, Singularity <i>Pipelines:</i> Nextflow, Snakemake, CWL <i>Job Scheduling:</i> SGE, SLURM, LSF
Genomics	<i>NGS Assays:</i> DNA-Seq, RNA-Seq, Methyl-Seq, ATAC-Seq, ChIP-Seq, HiC, Single-Cell (10X, Fluidigm) <i>Bioinformatics:</i> Samtools, BWA, GATK <i>Other:</i> 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 <i>A Big Data Approach to Learning the Type 2 Diabetes Regulome</i> Career transition award with 3 years tenure-track funding	2017 to Present
American Diabetes Association Postdoctoral Fellowship <i>A Multi-Tissue and Multi-Omics Investigaton of Type 2 Diabetes</i> Postdoctoral fellowship with up to 3 years salary and research support	2017 to Present
NIH Intramural Sequencing Center Pilot Grants <i>Four separate project proposals funded</i> Institutional award with funding for sequencing services	2014 to 2016
Department of Health and Human Services Ignite <i>LabGenius: The Smart Lab Notebook for Scientists</i> 3-month incubator program to fund innovative projects within HHS	2014
Dean's Distinguished Dissertation Award (Department Nominee) <i>University of North Carolina at Chapel Hill</i> Nominee from the Bioinformatics and Computational Biology program	2014
Verne Chapman Young Scientist Award <i>International Mammalian Genome Society</i> Best trainee talk at the International Mammalian Genome Conference	2013
Chicago Prize <i>Complex Traits Consortium</i> Best graduate student talk at the Complex Traits Consortium meeting	2013
Genome Research Award for Outstanding Poster <i>International Mammalian Genome Society</i> Outstanding poster at the International Mammalian Genome Conference	2010

Leadership & Service

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