

Ryan P Smith

linkedin.com/in/rpseq

github.com/rpseq

ryan.smith.p@gmail.com

+1-319-899-0190

ABOUT ME

I am passionate about learning and developing new technologies as well as tinkering with complex systems—especially genomes and software. After 7 years of genetics and bioinformatics training in academia, I am seeking industry experience with specific interests in data science and software engineering.

- **Relevant Interests:** Distributed computing, containers, automation, object-oriented and functional programming, data visualization, bioinformatics, genetics/genomics, DNA editing

EXPERIENCE

- **The McDonnell Genome Institute, Washington University** St Louis, MO
Graduate Research Scientist, Ira Hall Lab, Computational Genetics Feb 2015 – Present
 - **Distributed Computing, Bioinformatics:** Process and analyze population-level genome sequencing data to improve our understanding of the causes and consequences of structural variation in human genomes. Most work is done in a Linux environment with custom Bash pipelines, distributed across a Dockerized computing cluster. Pipelines often use bioinformatics tools combining Bayesian statistics and machine learning approaches.
 - **Single-cell Sequencing:** Develop novel computational and molecular biology methods to sequence the genomes of single mammalian neurons, in collaboration with the Scripps Research Institute at University of California, San Diego. Typically process up to 10 TB of raw Illumina sequencing data within a 36 hour period.
 - **Teaching Assistant:** Advise students in characterizing the genome sequences of unknown bacteriophages. Hands-on computer lab course using a series of bioinformatics tools to identify genes and predict their function, as well as classifying novel phages into existing phylogenetic groups.
- **University of Iowa** Iowa City, IA
Undergraduate Research Fellow, Adam Dupuy Lab, Cancer Genetics Aug 2010 – May 2014
 - **Genome Editing:** Designed methods for viral genetic engineering in mouse models of human cancers and a sequencing method for detecting resulting transgene insertions.
 - **Bioinformatics:** Processed high-throughput genome sequencing data using Linux bioinformatics tools followed by ad-hoc statistical analyses and data visualization in Python and R.

PROGRAMMING SKILLS

- **Languages:** Bash, Python, R (ggplot), awk, SQL, C++, L^AT_EX, Mathematica
- **Technologies:** Linux/Unix, OSX, sed, git, Docker, CircleCI, IBM Platform LSF, Oracle Grid Engine

EDUCATION

- **Washington University** St Louis, MO
PhD (In Progress), Molecular Genetics and Genomics Aug. 2014 – Present
- **University of Iowa** Iowa City, IA
BS, Microbiology and Informatics; GPA: 3.7 Aug. 2009 – May. 2014
 - **Relevant Coursework:** Intro to Computer Science, Programming for Informatics, Biostatistics, Programming with C++, Bioinformatics Techniques, Networking and Security, Human Computer Interaction, Database Management, Strategic Management of Technology, Informatics Capstone Project

PUBLICATIONS

- **Jennifer L. Hazen, Michael A. Duran, Ryan P. Smith*, Alberto R. Rodriguez, Greg S. Martin, Sergey Kupriyanov, Ira M. Hall, Kristin K. Baldwin** “Using Cloning to Amplify Neuronal Genomes for Whole-Genome Sequencing and Comprehensive Mutation Detection and Validation” *Genomic Mosaicism in Neurons and Other Cell Types. Neuromethods*, vol 131. September 2017
- **Ryan P. Smith*, Jesse D. Riordan, Charlotte R. Feddersen, Adam J. Dupuy** “A Hybrid Adenoviral Vector System Achieves Efficient Long-term Gene Expression in the Liver via PiggyBac Transposition” *Human Gene Therapy*, 26(6):377-85. June 2015
- **Jesse D. Riordan, Luke Drury, Ryan P. Smith*, Ben T. Brett, Adam J. Dupuy** “Sequencing Methods and Datasets to Improve Functional Interpretation of Sleeping Beauty Mutagenesis Screens” *BMC Genomics*, 15(1): 1150. December 2014