

# Keenan Graham

<http://keenangraham.github.io>

3165 Porter Drive  
Palo Alto, CA 94304  
(828) 280-5083  
ksgraham@stanford.edu

## EXPERIENCE

**Department of Genetics, Stanford University School of Medicine —  
Principal Investigator: Mike Cherry, Ph.D.**

May 2017 - Present

- **Develop software** and wrangle data at the **ENCODE Data Coordination Center** including:
  - *Backend.* Develop and maintain the ENCODE Portal's (<https://www.encodeproject.org>) Python/Pyramid application, **following proper Git workflows, Jira Sprints, unit and integration tests, and code reviews.**
  - *Backend.* **Troubleshoot failures** and upgrade deployment environment (Ubuntu, Postgres, Elasticsearch). Integrate database migration tool. **Write documentation for new features.**
  - *Frontend.* **Design and implement QANCODE**, an automated image-based comparison tool for regression testing in the browser.
  - *Pipelines.* **Modify genomic analysis pipeline accessioning code** to support wider variety of assay types. Introduce documentation tools to pipeline group.
  - *Pipelines.* Develop a general quality metric reporting tool that **performs complex data scraping and manipulation** and creates reproducibly formatted Google Sheets for use by consortium working groups.
  - *External scripts.* Maintain and update *pyencoded-tools*, scripts used by wranglers to submit and modify data, **in accordance with wranglers' specification.**
- **Design with lead data wranglers software solutions for modeling ENCODE objects**, controlling status, and synchronizing datastores in AWS and Google Cloud. Implement new features to significantly lower AWS storage costs.
- Act as resource for group members on backend permissions and views, metadata questions, Postgresql database inquiries, and DNAnexus failures. **Guide new group members in setting up development environment.**
- Analyze biosample, experiment, file and quality metric metadata.
- Present tutorial on JSON objects and interaction with the portal.

**Department of Microbiology and Parasitology, Genova Diagnostics —  
Laboratory Director: James Kelton, Ph.D.**

2015 - April 2017

- Write, revise, and execute custom automation scripts to extract data from the laboratory information system for use in research projects and

## TECHNICAL SKILLS

Data analysis and software development in Python and JavaScript

Custom dashboards with **D3.js**, HTML, CSS, Jupyter Notebooks, and Markdown

Database creation and **querying in SQL and Elasticsearch**

Web applications with Pyramid/Flask/Django, React

Git, Bash, AWS, Emacs

## LABORATORY SKILLS

**Diagnostic parasitology** and microscopy

Bacterial culture and antibiotic sensitivity assays

MALDI-TOF mass spectrometry

**Western blot**

PCR and DNA sequencing

## INTERESTS

Bioinformatics

Hierarchical **Bayesian analysis** and statistical modelling

Interactive data visualization

**Machine/deep learning**

laboratory monitoring including:

- All parasitology results from 2014 and 2015 (over 120,000 samples) and patient demographics (age, gender, location)
- All negative parasitology results since 2001
- All *Clostridium difficile* results since 2010
- Clean, transform, calculate summary statistics, and **perform complex analysis on data with Python** and Jupyter notebook.
- Create publication-quality visualizations and reports to share with supervisors and display in the laboratory.
- Implement classical statistical hypothesis tests and Bayesian analysis to answer epidemiological questions.
- Liaise with IT Department to create custom SQL queries of the backend database in order to monitor the daily input and output of the laboratory.
- **Build daily parasitology monitoring tools** and interactive, web-ready visualizations in D3.js, JavaScript, HTML, and CSS.

2013 - April 2017

- Examine via light microscopy trichrome-stained slides and iodine-stained concentrated wet mount preparations in order to **identify intestinal protozoa and parasitic helminths/ova in patient samples submitted for clinical diagnostics**.
- Complete >20 hours of continuing education per year and participate quarterly in College of American Pathologists proficiency testing and in-house quality assurance.
- Cross-train in bacteriology methods of sterile plating, Vitek 2 identification and antibiotic sensitivity assays, and MALDI-TOF mass spectrometry.
- **Have resulted more than 20,000 patient samples.**

**Department of Clinical Genetics, Fullerton Genetics Center, Mission Hospital — *Laboratory Director: Jack Tarleton, Ph.D., FACMG***

2013 (Volunteer Project)

- Identify novel mutations in the *CLCN1* gene leading to myotonia congenita.
- Converse regularly with the laboratory director on general topics of clinical genetics and NGS technology and specific topics of project-related data and goals.
- **Correct inconsistencies between lookup table of point mutations and raw DNA exon sequences of the gene.**

**Department of Epidemiology, UNC Chapel Hill Gillings School of Global Public Health — *Principal Investigator: Carla Cerami Hand, M.D./Ph.D.***

2010 - 2011

- Determine parasitemia in blood samples infected with the malaria parasite *Plasmodium falciparum*.
- Isolate the non-erythropoietic, tissue-protective erythropoietin heteroreceptor.
- Collaborate and attend laboratory meetings with practicing research scientists, Ph.D. candidates, and laboratory technicians.
- **Senior research thesis based on experiments performed in this laboratory.**

## EDUCATION

**University of North Carolina at Chapel Hill, Chapel Hill, NC —**  
*Bachelor of Science in Biology*

2011

Coursework in the Biology of Blood Diseases, Molecular Biology, Genetics, Organic Chemistry, Biochemistry, Modern Physics, Ordinary Differential Equations, Linear Algebra, Advanced Mathematical Methods.

## RECENT SOFTWARE PROJECTS

**HapPyNet —** *Haplotype-block-based dimensionality reduction for complex variant-disease associations.*

NCBI-UCSC Genomics Hackathon, Santa Cruz, CA, April 2-4, 2018.

**More at** <https://github.com/NCBI-Hackathons/HapPyNet>

**Biosynthetic Gene Cluster Data and k-Mer Search Portal —**  
*Applying k-mers to find similar DNA sequences in the Minimum Information about a Biosynthetic Gene cluster (MIBiG) data repository.*

October 2016.

**More at** <https://biosynthetic-kmers.herokuapp.com>

**The Epidemiology of Intestinal Parasites —** *The Prevalence of Protozoal Infection in 120,000+ Clinical Fecal Samples, United States, United Kingdom, Northern Europe, Canada, Australia.*

Genova Diagnostics, Asheville, NC, 2014- 2015.

**More at** <http://keenangraham.github.io/project/project.html>