Python Hackathon

UCLA QCBio Collaboratory

Thiago Mosqueiro

Renaud Dessalles

Simon Mitchell

Our goal for today

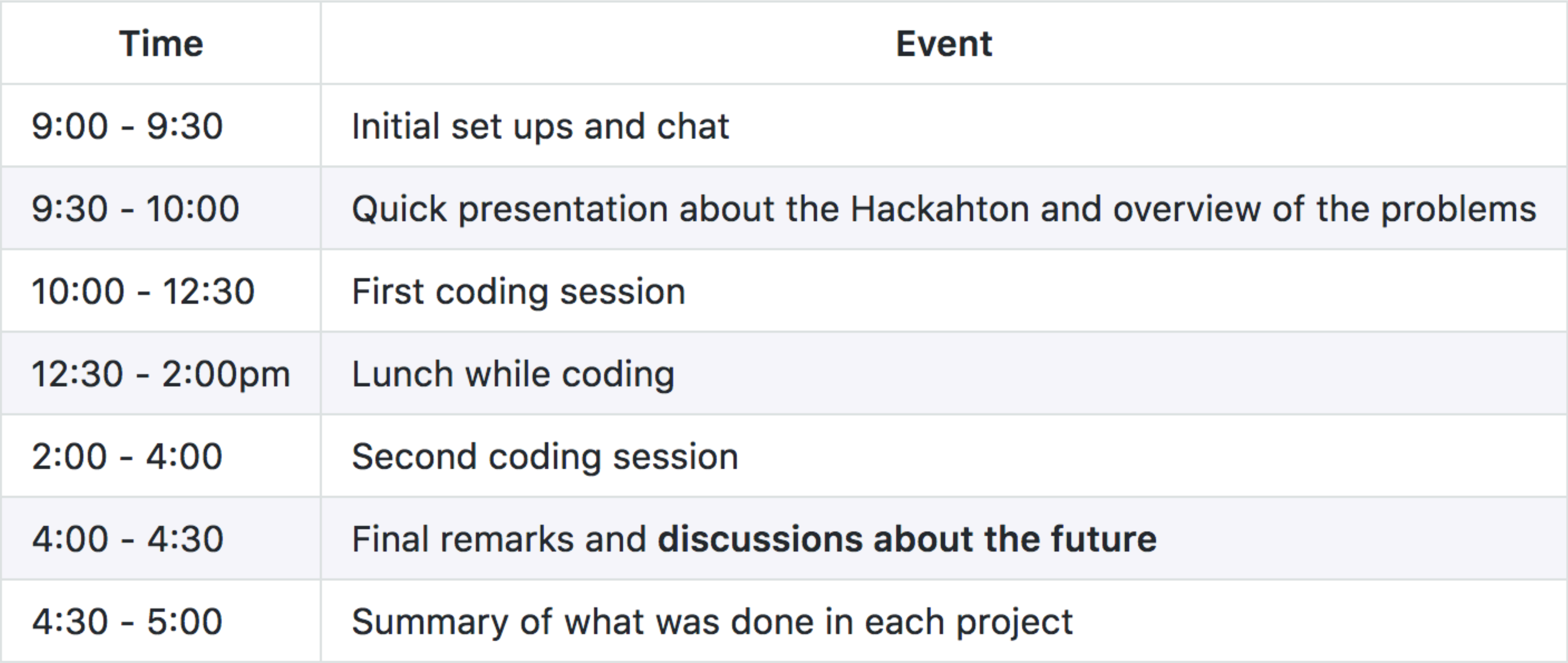
- Use this time to code together and exchange ideas and experiences among the participants

- Learn more about how Python is used in real-life projects applied to bio-related fields

- Collectively solve a set of projects of interest to the QCBio community

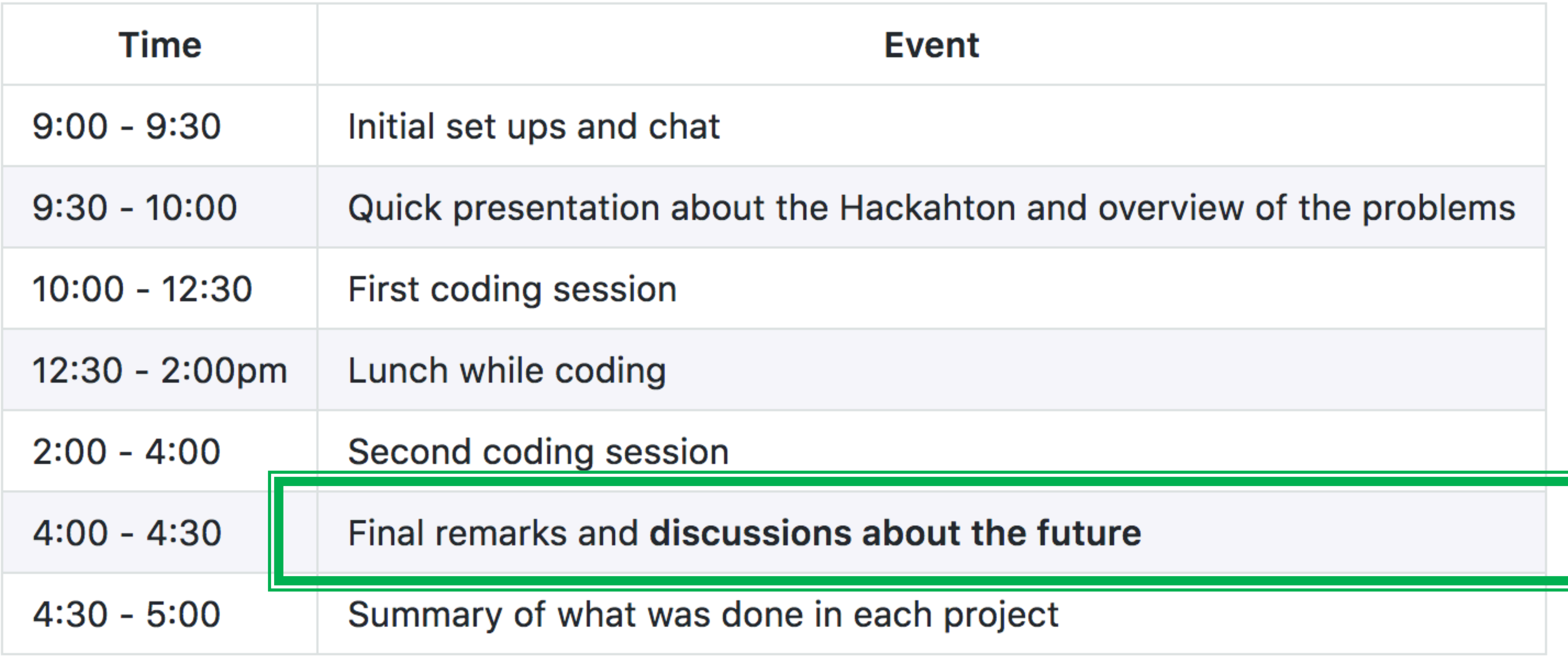
- Collaboratory fellows will be available to help during the development of the projects

Schedule for today



Time	Event
9:00 - 9:30	Initial set ups and chat
9:30 - 10:00	Quick presentation about the Hackahton and overview of the problems
10:00 - 12:30	First coding session
12:30 - 2:00pm	Lunch while coding
2:00 - 4:00	Second coding session
4:00 - 4:30	Final remarks and **discussions about the future**
4:30 - 5:00	Summary of what was done in each project

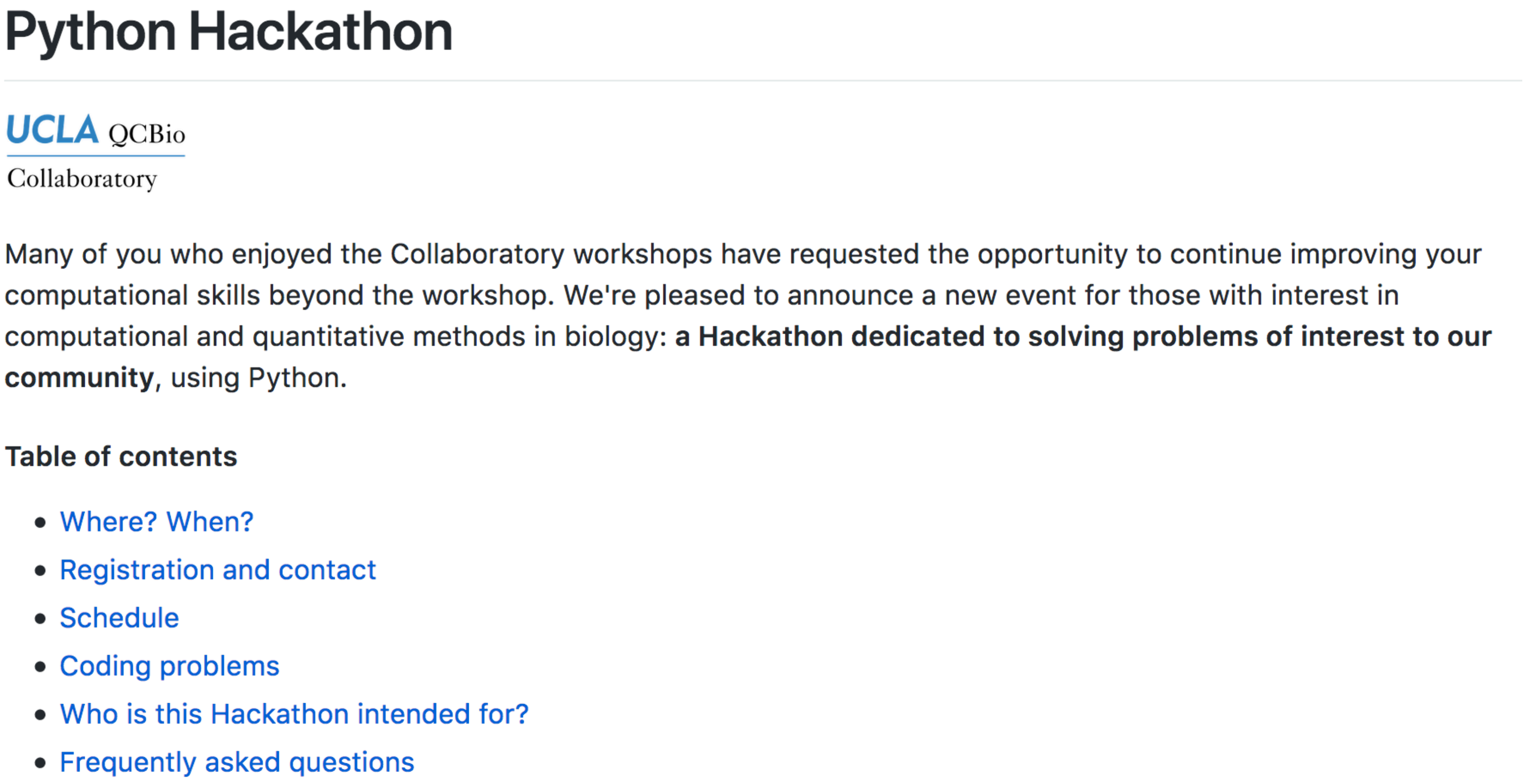
Schedule for today



Time	Event
9:00 - 9:30	Initial set ups and chat
9:30 - 10:00	Quick presentation about the Hackahton and overview of the problems
10:00 - 12:30	First coding session
12:30 - 2:00pm	Lunch while coding
2:00 - 4:00	Second coding session
4:00 - 4:30	Final remarks and **discussions about the future**
4:30 - 5:00	Summary of what was done in each project

The material is currently available on
https://github.com/thmosqueiro/UCLA-Collaboratory_Hackathon/

or https://goo.gl/c4LtW6



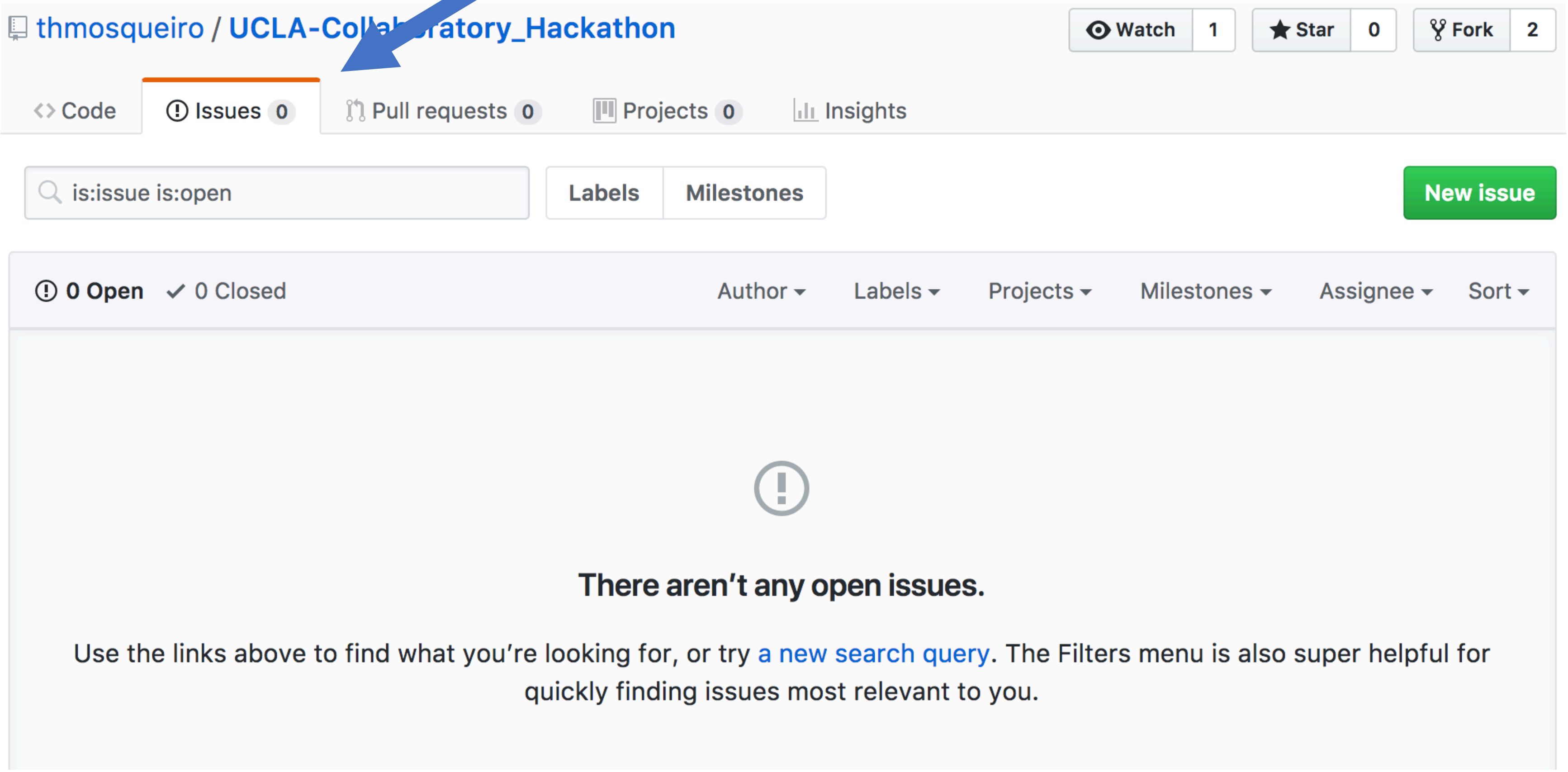
Python Hackathon

UCLA QCBio
Collaboratory

Many of you who enjoyed the Collaboratory workshops have requested the opportunity to continue improving your computational skills beyond the workshop. We're pleased to announce a new event for those with interest in computational and quantitative methods in biology: **a Hackathon dedicated to solving problems of interest to our community**, using Python.

Table of contents

If you have some material you want to share with everyone else, use the Issues page on the GitHub.



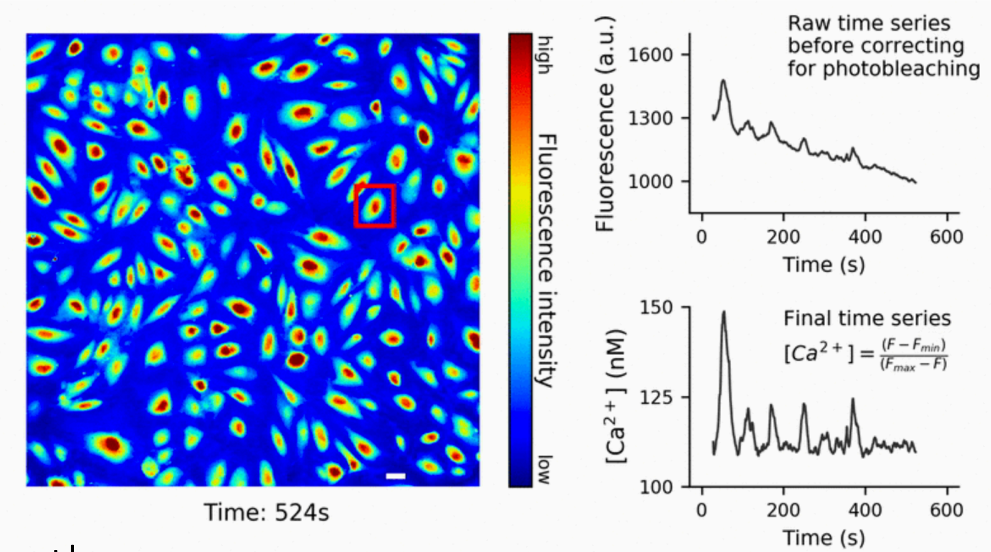
There aren't any open issues.

Use the links above to find what you're looking for, or try a new search query. The Filters menu is also super helpful for quickly finding issues most relevant to you.

Projects

1- Analysis of calcium imaging

- **Goal:**
  A Jupyter notebook that summarizes how we extracted calcium time series

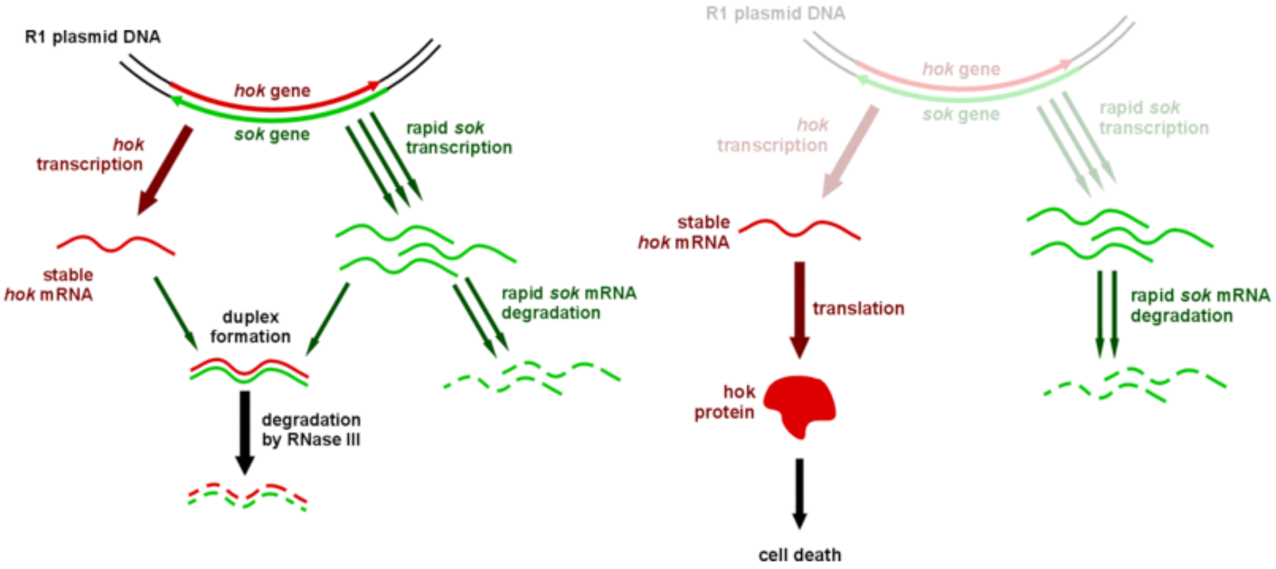


- **Technical Challenges:**
  - Handling images and videos with python
  - Dealing with photobleaching and estimating calcium concentration
  - Applying regression on a set of time series
  - Extracting statistics based on a set of cells

- **Dataset:**
  - We will use a dataset used in a recent publication by Julia Mack @ Arispe Lab

2- Gene expression with Gillespie algorithm

- **Goal:**
  Using the Hok/Sok system as an example, we will explore the Gillespie algorithm
- **Technical Challenges:**
  - Create a model that represents a real biological system.
  - Learn how to implement the Gillespie algorithm
  - Determine the parameters for the simulation

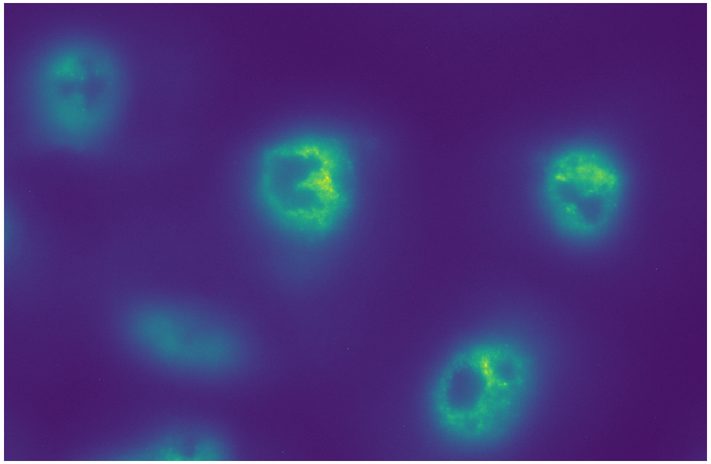


3- Automation of pipeline

- **Goal:**
  Automate a pipeline for image processing using efficient use of parallel resources and memory management.

- **Technical Challenges:**
  - Basics of parallel processing in Python
  - Automating the analysis of batches of large images
  - Good practices for a memory efficient implementation of parallel processing in Python

- **Dataset:**
  - We will use a dataset currently being developed by Rob Foreman @ Wollman Lab

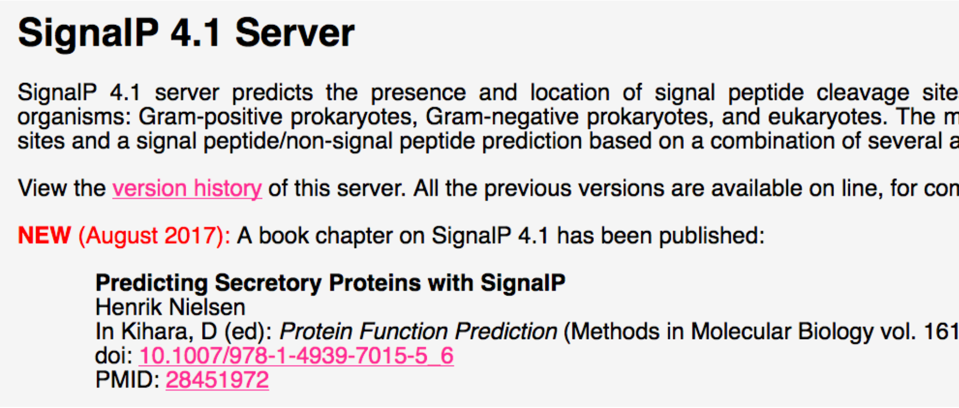


4- Automated job submissions

- Goal:
  Write a script that prepares and sends sequences to SignalP server, an online service that predicts cleavage sites of signal peptides.

- Technical Challenges:
  - Automating genomics analyses based on online applications
  - Dealing with a bottleneck when analyzing a large dataset
  - Learning how to construct web crawlers

- Dataset:
  - Chlamydomonas reinhardtii's
  - proteome, availablne on the Genome Portal hosted by the Joint Genome Institute



SignalP 4.1 Server

SignalP 4.1 server predicts the presence and location of signal peptide cleavage site
organisms: Gram-positive prokaryotes, Gram-negative prokaryotes, and eukaryotes. The m
sites and a signal peptide/non-signal peptide prediction based on a combination of several a

View the version history of this server. All the previous versions are available on line, for con

NEW (August 2017): A book chapter on SignalP 4.1 has been published:

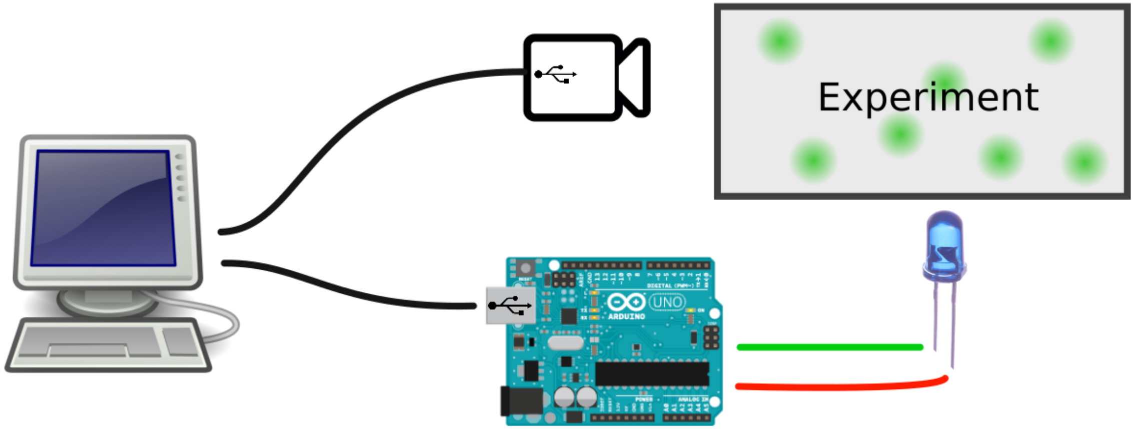
Predicting Secretory Proteins with SignalP
Henrik Nielsen
In Kihara, D (ed): *Protein Function Prediction* (Methods in Molecular Biology vol. 161
doi: 10.1007/978-1-4939-7015-5_6
PMID: 28451972

5- Integrating camera and arduino

- **Goal:**
  Construct an acquisition system that integrates a camera with an LED array controled by an arduino.

- **Technical Challenges:**
  - Use Python to synchronize an LED array and a camera
  - Writing very simple code for arduino
  - Control a camera from Python

- **Dataset:**
  - We will use the camera and arduino to capture data during the Hackathon



Let’s get started!