Daily Schedule for March 2019 Santa Cruz Pangenomics Hackathon!

# Directions to Seymour Center <Room>: (HQ):

# Contact Information

Organizers: Benedict/Tina/Alexis/Kelly from UCSC and Ben Busby (Ben.busby@gmail.com)

If you get lost or waylaid for any reason, Ben’s cell is 410-868-1073

Ben doesn’t always pick it up right away, so feel free to blow it up with text messages or whatever, or just send an email/slack message.

# Slack Channel

*https://join.slack.com/t/pangenome39/shared\_invite/enQtNTY1NDUzNTYyMTE1LTJlOGFkNDNiZDhkODgyZjE1MDRiOTViMzZjM2RmM2U2YWQzOTQ4ODYzZDQxOGRkYzg1OWMwMWExZmM0NDg0YTY*

# Hotels

Kelly was able to secure a great rate from the Marriott located in walking distance to the Hackathon site. Participants can call and use code UCSC Genomics Institute to secure the block rate for the Hackathon.

3/24 rate $139

3/25 rate $155

3/26 rate $155

For directions to the hotel, Seymour Center = (Hackathon Site), airports, etc. go to the map link on the hotel website very user friendly directions link.

<https://www.marriott.com/hotels/travel/sjccr-fairfield-inn-and-suites-santa-cruz/>

**\*\*Block rate/Cutoff date 3/7/2019\*\***

Please encourage everyone to make reservations before this date to secure everyone stays at the same hotel.

**Please note that this is a dynamic document, and therefore subject to change at the discretion of the organizers and participants.**

*Italicized text: Optional Items*

# Projects

## Construction

1. **Haplotype Definition Consortium!** -- Leads:
   1. <Github repo/folder>
   2. <Cloud login info>
   3. <Documentation link>
2. **Building haplotypes and graphs from full *de novo* assemblies!** -- Leads: Paten,
   1. <Cloud login info>
3. **Tools that could work on the pangenome and new data types!** -- Leads: Marschall
   1. <Cloud login info>
4. **Truth sets and validation** -- Leads: Erik Garrison, Justins from NIST
   1. <Cloud login info>

## Application

1. **RNA-seq assortment** -- Lead: Busby, ?
   1. <Cloud login info>
2. **Using graphs to infer structural variants from short reads** -- Lead: Sedlazeck, ?
   1. <Cloud login info>
3. **Visualization and human interfaces to the pangenome** -- Lead: TBD (Possibly Maria Nattestad, Greg Schuler from NCBI?)
   1. <Cloud login info>
4. **“Sanity Check” team** -- Leads: Paten,
   1. <Cloud login info>

Draft of manuscript in F1000R format:

## <Ben will provide a link>

## GitHub

### Sample github documentation!

<https://github.com/NCBI-Hackathons/GoodDoc/blob/master/README.md>

<https://github.com/NCBI-Hackathons/PubRunner>

<https://github.com/NCBI-Hackathons/MetagenomicAntibioticResistance>

<https://github.com/NCBI-Hackathons/Staph_aureus_viewer>

**PLEASE ASSOCIATE KEYWORDS WITH YOUR GITHUB REPOS!**

**Consider registering them with osf.io or RRIDs**

**Markdown cheatsheet**: https://github.com/adam-p/markdown-here/wiki/Markdown-Cheatsheet

## Manuscripts

1. (1) Introductory Paragraph
2. (2) Methods Paragraphs
3. (2) Results Paragraphs
4. (1) Discussion Paragraph

Please edit your manuscripts!

Send figures for the manuscript to Lisa at lisa.federer@nih.gov

### Sample manuscripts!

PubRunner: <https://f1000research.com/articles/6-612/v1>

DangerTrack: <https://f1000research.com/articles/6-443/v1>

# Timeline

## Wednesday, March 20th (Online)

**Human Pangenomics Data Upload Party!**

Wed, Mar 20, 2019 12:00 PM - 8:00 PM EDT

**Please join my meeting from your computer, tablet or smartphone.**

<https://global.gotomeeting.com/join/555721317>

**You can also dial in using your phone.**

United States: +1 (872) 240-3311

**Access Code: 555-721-317**

## Monday, March 25th

*9:00: Coffee and Breakfast (Provided)*

9:15-9:45: Introductions, Administrative Announcements and ***Building Functional Software and Data Indices***

* + Discussion of Roles
    - **Lead**
      * Understands the science and assigns roles
    - **Sysadmin**
      * Decides where data and tools go, and implements git for versioning.
      * Deals (in conjunction with technical assistance) with any technical issues (team tech lead)
    - **Writer**
      * Details both product and process for:
        + Documentation
        + Manuscript
    - Discussion of Output
    - GitHub
      * Live cloning of repos
      * Docker/Singularity
    - Publishing
      * Advertising and social media
      * RRID registration
* 9:45 - 10:00: Role assignments in teams
* 10:00 - 12:15: Sysadmins and Data Acquisition get data, if necessary
  + Everyone else work on first sketch of plan (one page google doc, or similar)
* 1:00 - 1:30: Working lunch (Provided).
* 1:30 - 6:30: Pipeline building <script building>
* *6:30 - 8:30: DINNER (on your own or with group members). Either local restaurant or in <Breakout Room>*
* *7:30 - 8:30: Implement any remaining code/start pipelines (remote is an option) <script building>*

## Tuesday, March 26th

* *8:30: Grab Coffee/Breakfast (Provided)*
* 9:00 - 9:15: Keynote <TBD>
* 9:15 - 12:30: Data cleanup, code fixing, and presentation generation
* 12:30 - 1:30 Working lunch. Each section presents three to four slides; opportunity for exchange of ideas (*and personnel if necessary -- group lead solicitation*). *Lunch in <room?>*
* 1:30 - 6:00: Pipeline fixing and implementing suggestions from other group; more data runs
  + 3:00 - 4:00 Writer’s Group break off 1
* *6:30 - 8:30: GROUP DINNER (provided) opportunity for feedback on organizational structure and future projects.*

## Wednesday, March 27th

* *8:30: Breakfast provided.* *More opportunity for feedback on organizational structure and future projects*
* *9:00 am Bioinformatics Software Repos?*
* 9:15 - 12:00: Finishing pipelines
* 11:40 - 12:10: Working lunch
  + *Lunch in <Room?>.*
* 3:00 - 3:30:
  + Writers break off for writers group
  + Editorial representative from NLM will be there to assist.
  + Everyone else documents code and writes READMEs
    - Pipeline testing!
      * Bug fixes
    - Presentation generation
* 3:30 - 3:45: Marketing, community interaction, and longevity of code -- TBD
  + Who is going to deal with future pull requests?
  + Awards ceremony!
* 3:45 - 4:30: Final Presentations (from Github repos) (Open to Scientific Public)