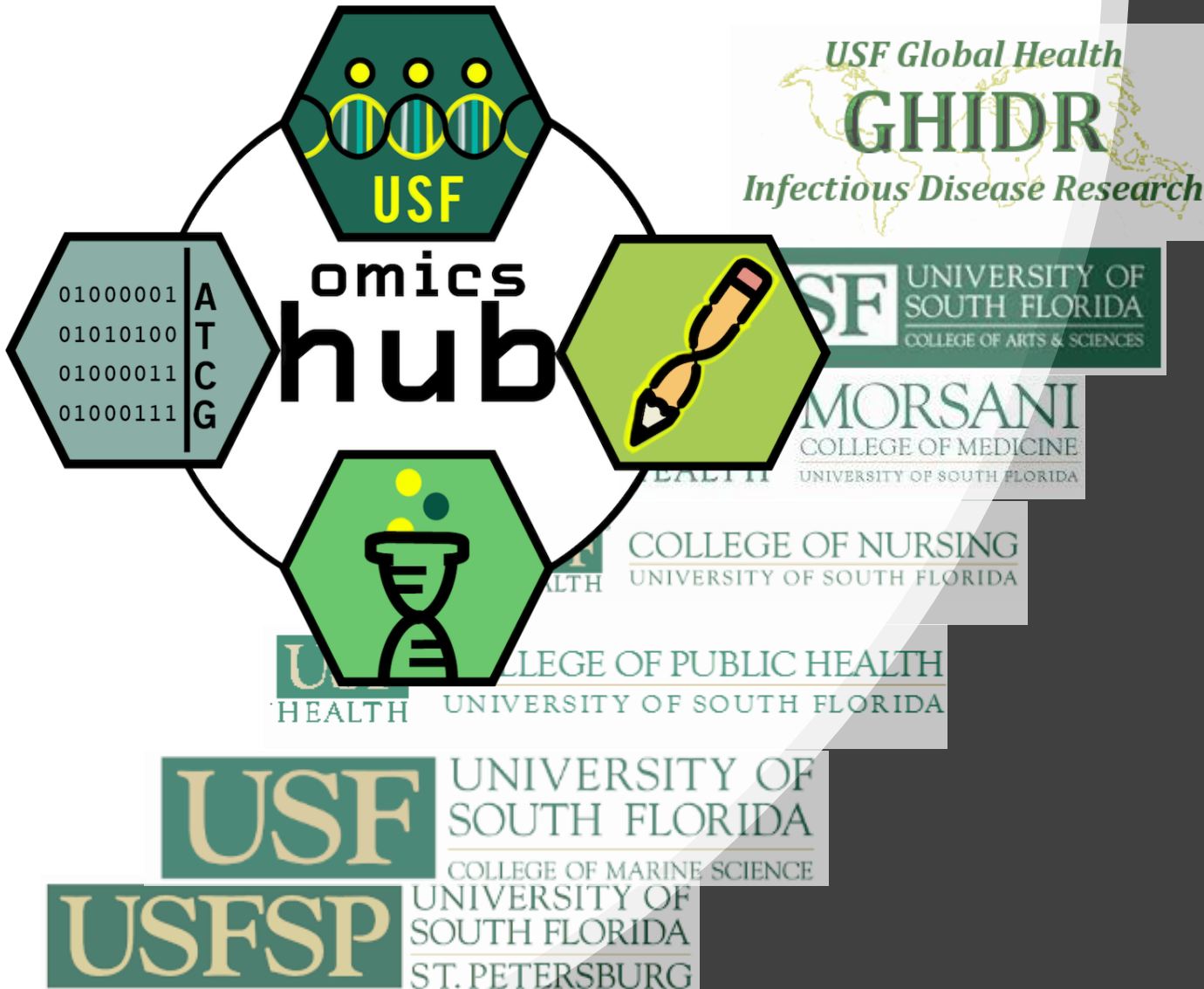




UNIVERSITY OF  
SOUTH FLORIDA  
A PREEMINENT RESEARCH UNIVERSITY

GENOMICS  
PROGRAM



# Microbiome Data-Analysis Workshop

**Jenna Oberstaller, PhD**

Interdisciplinary Sciences Liaison, USF Genomics

Managing Consultant, USF Omics Hub



```
graph TD; A[USF Genomics, the Hub] --> B[Microbiome-research at USF]; B --> C[Workshop purpose and goals]; C --> D[Resources]
```

USF Genomics, the Hub

Microbiome-research at USF

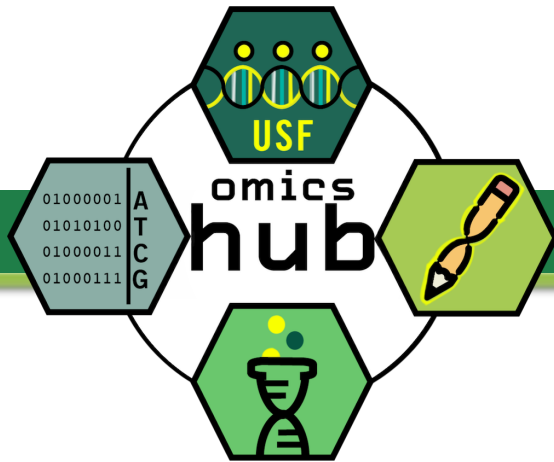
Workshop purpose and  
goals

Resources

ORIENTATION

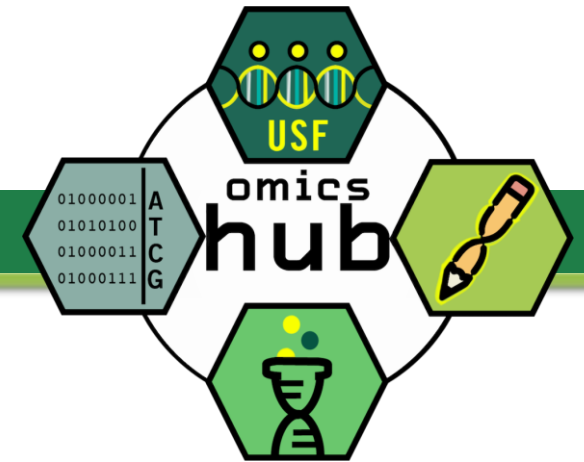
# PROGRAM STRUCTURE AND MISSION OVERVIEW

## USF Genomics Program



- ❖ Establish a collaborative and open community to support all areas of Omics research at USF

### Omics Hub



- ❖ Build awareness for in-house capabilities
- ❖ promote departmental and community growth
- ❖ Computational and laboratory consulting in partnership with the Core
  - ❖ Encourage grants submission as collaborators
  - ❖ Provide training (bench, data analysis)

## USF Genomics Core Facility

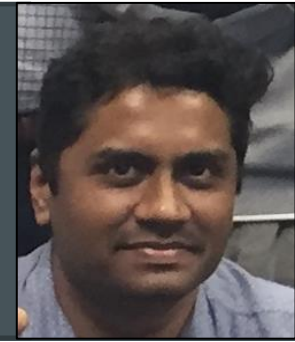
- ❖ Core equipment to which certified users have **SUPERVISED** access after training and project-consultation with our Core experts

# USF MICROBIOME INITIATIVE

- CROSS-USF and beyond
- Build USF's profile as a Leader in Microbiome-Research
- Support USF Researchers and foster interdisciplinary collaborations
  - Microbiome-awards

Career biology  
experts here  
with you to

Featuring special guest-star Anujit Sarkar, PhD  
Alman/Groer labs  
Microbiome-specialist



## Computational expertise

genomics@usf.edu



Jenna Oberstaller, PhD



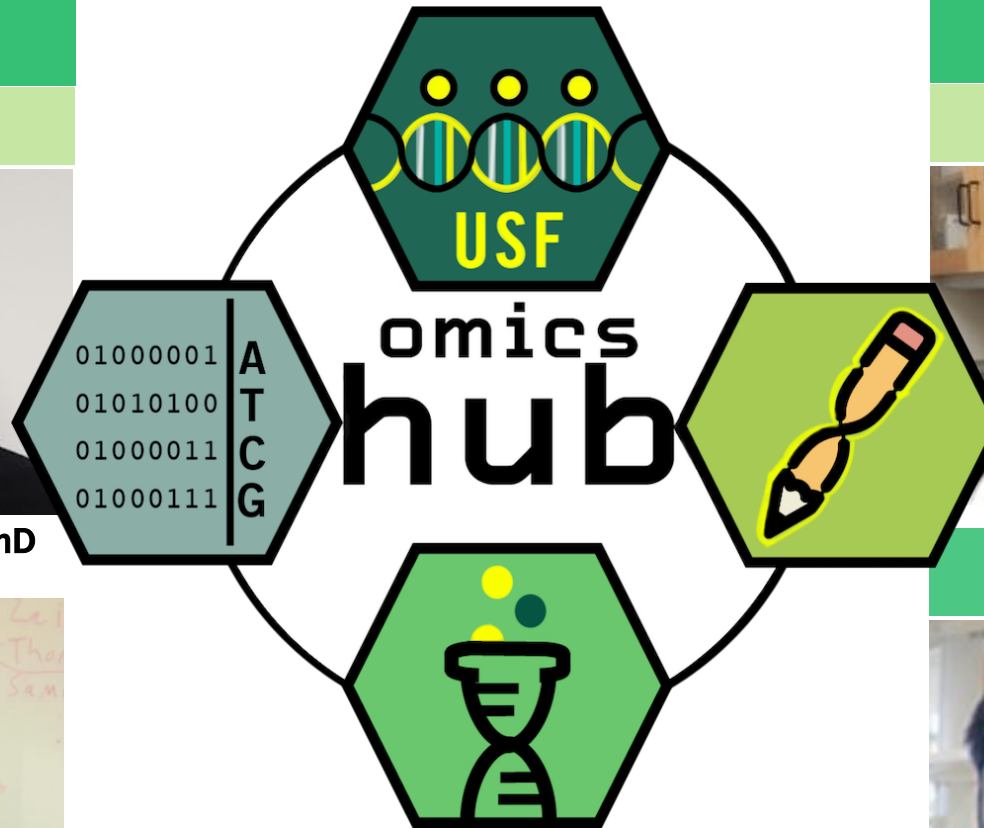
Charley Wang, PhD



Thomas Keller, PhD

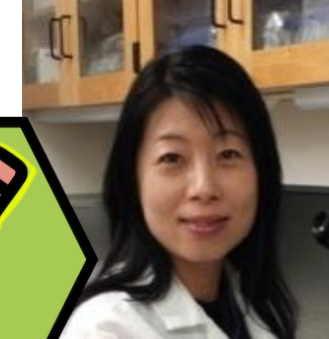


Justin Gibbons, PhD



## Laboratory expertise

genomics@usf.edu



Min Zhang, MD  
Core instrumentation  
expert

## Honorary Hub affiliate



Swamy Rakesh Adapa, MS  
Founding Hub member,  
Jiang-lab associate and  
Renaissance Man

genomics@usf.edu

# WORKSHOP STRUCTURE

## CRASH COURSE!

Introduction to  
Microbiome Data Analysis

Computational methods +  
best practices for  
experimental design

**Wednesday**

## Hands-on training

Taxonomic analysis  
Experimental Design

Taxonomic analysis  
OTU/ASV Generation

**Thursday**

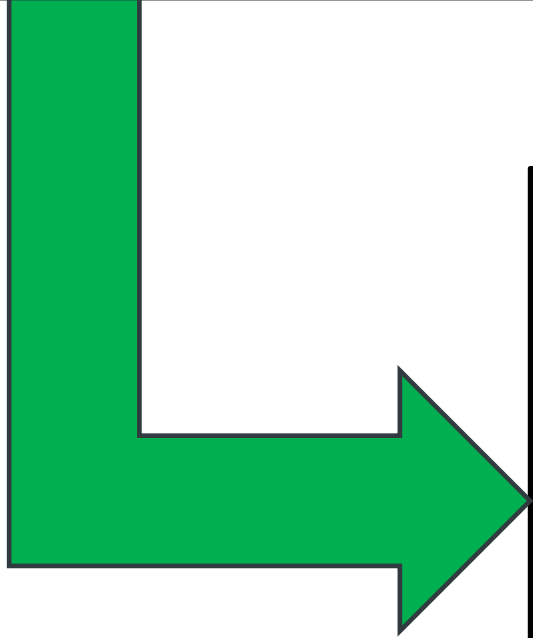
Taxonomic analysis  
Visualization

Tataxonomic analysis  
Visualization

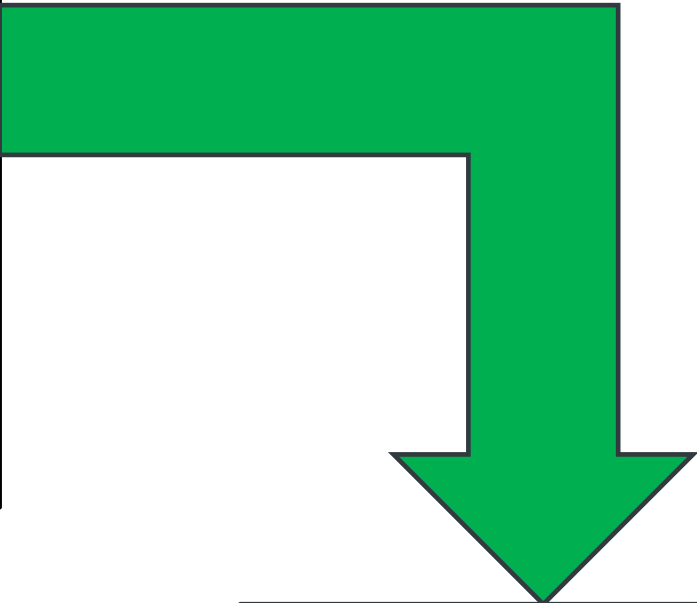
**Friday**

*Prior sequence analysis or command line experience are **not** required*

**Experimental question**



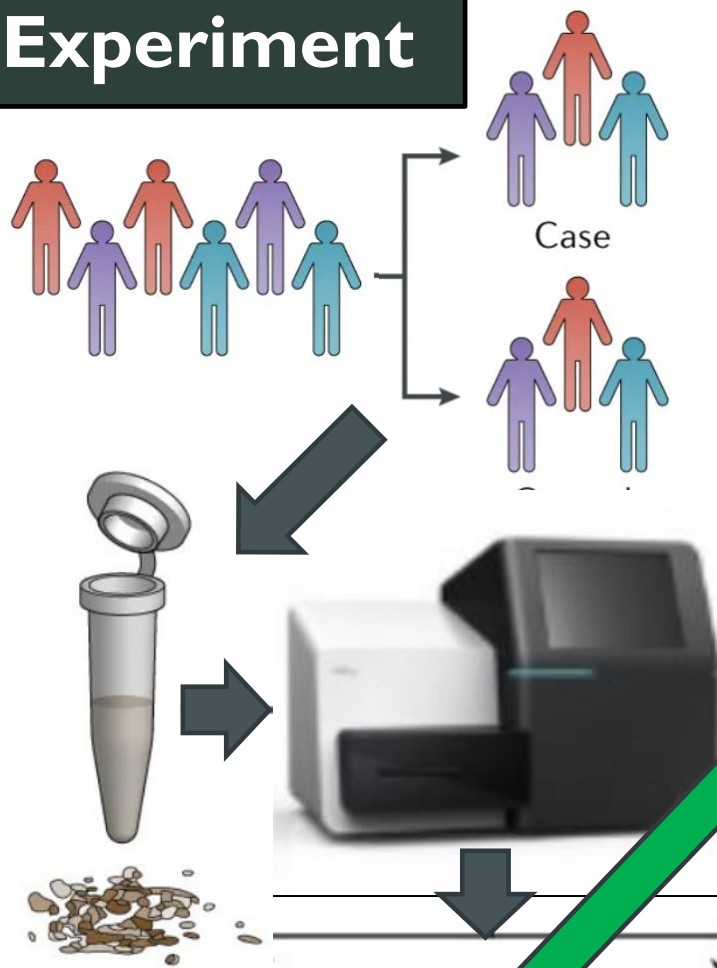
**Omics MAGIC**  
**?????**



**BIOLOGY**

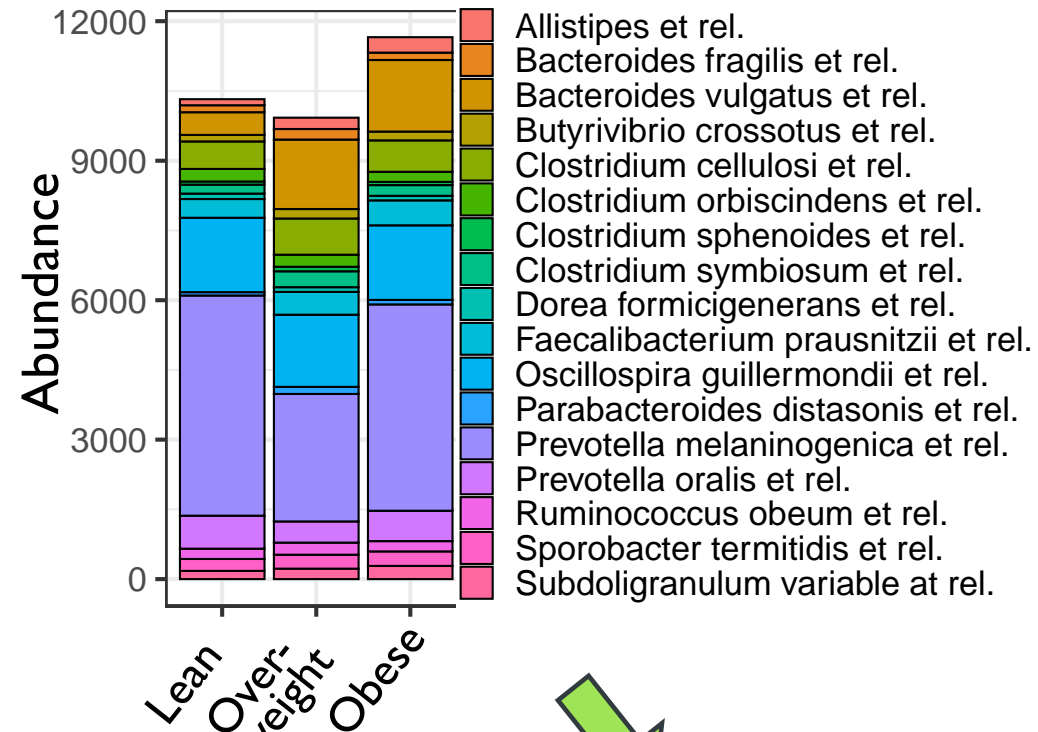


# Experiment



```
>fasta sequence 1
CCGcAagCGATtCGTAGCGCGCTGGaCGTCTACGGCGAGAACGTccGCAAGATCGTGCCG
ATCGACAGCCCGcAGTTCCGGCCGCGTCACCTGCGTctGCATCGGCGCGATGATGGTCGGT
AGCATCCACATGACCGTCGAGGAGGGCCAGCAGGTGAAGCGCGCCAGgAGTTTGGGTAC
TTCGCTTTCGGTGGCTCGACGATCGTCTGCTCTTCGAGAAGGGCGTCGTGCAGTGGGAC
GAGGACTTGCTGATCAACAGCGCGCGTGCCTCGAGACTCTCGTGCGCGTCGGCATGGGT
ATCGGgcAcGGCAaCGcAAGCCGAACACGGACGGTAGTAGTGACAGTAGACGTGGAT
cGACcGCCTTtcTGGtGTGCATGTAGATACCTCGCGGAGTAGCCTACCcGGACCCTCCTA
ATAACGGACTTGTATGCAAACGTCCGCGCGGGTCCGCGACTACGTGCGCGCGTGGCCAG
>fasta sequence 2
GACCCcTCCCTtATCCTCTCCGCGTTCATCACTctcGGTCCCCTCGGCCAAGGTGGCTTC
TCGCTTGTCAACGGCCAGGACCTCTCTGAGATCCTGCCGCTGCATATGGGcGACGAG
TTCCTTGTAGCTCTCGCGGGCCAGATGGTCTTCGCGCGGTGCTTCTGGGGCGCCTTC
TTCCTTCTCCCTTGGCGTCGCTGGATTCTGCTGGCGATCATCTCCATCGGACACGTG
GCGCAGCTCACAGATTCTTCTCGATGGCGTACTGGGCGCTTATTTCCCGAACGGG
TGTTCGCCCTCCTATGCGTGCagTCGCGAAGGTGCTGGACAGTCTGTCTTCAGGGCG
AtcGGtGCTGCGTGGACCGGGGTGTTTTCTGCTCTGGATCtctGTCTTCACTCGCAGC
ATACCGTCGTTTCATGATGGCTCCATGTTCAAGGCGCCTtATGTTCCAGACCTTGTGAC
CCCGCAAAGGCAAAACATCGATGAGGAAAAGGGTTTCAACAACCCACCCAGCAACAAC
```

## Taxon abundance by sample



	Sample1	Sample2	Sample3
ASV_1	29	284	325476
ASV_2	16851	17521	34
ASV_3	0	0	0
ASV_4	26566	17	0
ASV_5	23225	19	0

genomes  
linking phylogeny & function,  
genome diversification

**BIOLOGY**



# CONCEPTS

- EXPERIMENTAL DESIGN is paramount
- BEST PRACTICES
  - at the bench
  - at the computer
- Data-analyses
  - QUALITY CONTROL
  - APPROPRIATE analyses to apply to answer YOUR biological question
    - many different methods, stats to be considered
- Tutorials for biologists
  - we can't remove ALL the code
  - but focus = straightest path from data to biology/interpretation

# HUB MICROBIOME WORKSHOP GITHUB PAGE

- Will go live tomorrow!
- Workshop-materials
- Walk-throughs
- Resources

Special thanks: Janelle Donglasan



## Microbiome Data-Analysis Workshop

### GOAL

In addition to understanding the importance of experimental design, we will walk through turning raw sequence data into useful counts data that we can use to visualize microbiome sample composition then discuss methods to extrapolate function from abundance data and ultimately arrive at biological insight.

### Table of Contents

1. Pre course Materials
2. Day One
  - >> Presentation Slides
3. Day Two
  - >> Presentation Slides
  - >> Generating ASV Tables from Microbiome-sample Sequencing Data
4. Day Three
  - >> Presentation Slides
  - >> Microbiome Data Visualization

### Pre course Materials

● [Best Practices for Analyzing Microbiomes](#) - This article discusses how all stages of conducting a microbiome study, from designing the experiment to collecting and storing the samples to obtaining insight from graphical displays of the sequence data, can substantially impact the result.

● If you do not have R and RStudio installed already, you can follow these [instructions](#) for both Mac and Windows. If you already have RStudio, make sure the latest R version is 4.0.1 -- "See Things Now" by clicking on [Global Options...](#) in the [Tools](#) tab. The version is also stated in the first line of the console when you first open RStudio. Follow the previous link to get the new version then uninstall the old version from your computer.

You can copy the library with your old packages from the previous version to the new version then update the packages by clicking [Check for Package Updates...](#) also in the [Tools](#) tab within RStudio.

If you have Windows, then a very easy way to update your R version and packages is by simply running the following code in RStudio console:

```
install.packages("installr")  
  
library(installr)
```

# USF Genomics home page



## USF Genomics

Home

Researchers

Grants

Degree Programs >

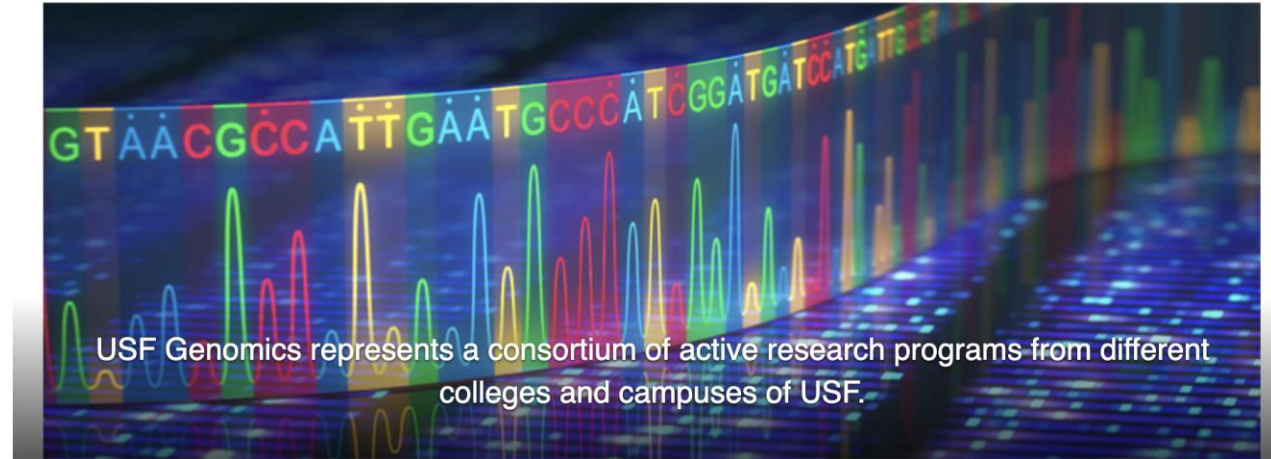
Undergraduate Opportunities >

Annual Symposium

College of Arts and Sciences  
News

Genomics Training

Request Services



## About USF Genomics

Our projects integrate modern genomic approaches to understand and help develop solutions for some of the major challenges affecting global health and especially our Florida community. USF Genomics research projects are naturally interdisciplinary with studies that cross many boundaries of traditional scientific fields, integrating bioscience approaches in experimental and environmental research with computation biology. Opportunities exist for new graduate students to join well-funded research programs to study disease causes from underlying drug resistance and virulence genes in pathogenic organisms to changes of microbial communities within our bodies to the marine environment.

Register and Submit Abstracts for the USF Genomics  
Annual Symposium on November 15, 2019

## Upcoming Events

USF Genomics Seminars: Tuesdays once a month,  
4:00-5:00 pm - IDRB 302

USF Genomics Seminar - Dr. Maureen Groer -  
Microbiome State of the Science - Tuesday,  
September 17, 4:00-5:00 pm - IDRB 302

USF Genomics Forum Discussion Series - Tuesdays  
once a Month, 5:00-6:00 pm - IDRB 302

USF Genomics Annual Symposium: Personal  
Genomics - November 15, 2019

Genomics Laboratory and Data-Analysis Training  
Courses: February, May, September 2019

# USF GENOMICS PROGRAM



**Your first contact for all  
Genomics Program inquiries**



**genomics@usf.edu**