**Syllabus for Bioinformatics Bootcamp (Fall 2020)**

**Dates and times:** Tuesdays 5-6:30PM, September 1 – December 1

**Location**: <https://zoom.us/j/97122219816?pwd=eGI5MTF6bzhPNUVaY1lPS1J5aFJpUT09>

**Topic:** Programming Fundamentals and RNA-Seq analysis

**Course Instructors:**

**Name**: Henry Miller, M.S.   
**Title:** 3rd year graduate student, IBMS PhD Program  
**Email:** [millerh1@uthscsa.edu](mailto:millerh1@uthscsa.edu)

**Zoom office hours:** 4-5PM Monday and 5-7PM Thursday

**Link:** <https://zoom.us/j/94153810537?pwd=YXlOalVyN1E3NmtFZFdGUk9NY1lVQT09>

**Name**: Paul Ramirez, M.S.  
**Title:** 3rd year graduate student, IBMS PhD Student  
**Email:** [ramirezp1@uthscsa.edu](mailto:ramirezp1@uthscsa.edu)

**Zoom office hours:** 5-7PM Monday

**Link:** <https://zoom.us/j/98595842069?pwd=a2h1R24yb01ocnBJaEg5RnBZRUtYUT09>

**Name**: Muku Bhandari   
**Title:** 2nd year graduate student, IBMS PhD Program  
**Email:** [bhandarim@uthscsa.edu](mailto:bhandarim@uthscsa.edu)

**Zoom office hours:** 5-7PM Wednesday

**Link**: <https://zoom.us/j/92530455838?pwd=TnRSRjJSVmtJa1FlR01zWTd2N1RBQT09>

**Name:** Yidong Chen, Ph.D.  
**Title:** Professor, Population Health Sciences  
**Email:** [cheny8@uthscsa.edu](mailto:cheny8@uthscsa.edu)

**Name:** Evelien Bunnik, Ph.D.  
**Title:** Assistant Professor, Microbiology, Immunology, and Molecular Genetics  
**Email:** [bunnik@uthscsa.edu](mailto:bunnik@uthscsa.edu)

**Name:** Bernard Fongang, Ph.D.  
**Title:** Assistant Professor, Biggs Alzheimer's Institute  
**Email:** [cheny8@uthscsa.edu](mailto:cheny8@uthscsa.edu)

**Name:** Siyuan Zheng, Ph.D.  
**Title:** Assistant Professor, Population Health Sciences   
**Email:** [zhengs3@uthscsa.edu](mailto:zhengs3@uthscsa.edu)

**Name:** Chris Chiu, Ph.D.  
**Title:** Post-Doc, Population Health Sciences  
**Email:** [ChiuY@uthscsa.edu](mailto:ChiuY@uthscsa.edu)

**Description and Learning Outcomes:**

This workshop will teach fundamental programming techniques with R and basic RNA-sequencing analysis approaches, such as differential expression analysis and gene heatmaps. Most sessions will consist of a 30-40-minute lecture, followed by hands-on activities guided by the instructors which, if not completed before the workshop ends, will be finished for homework. In the final two weeks, students will complete a final project which combines skills learned throughout the workshop.

Skills learned:

1. R coding basics –
   1. Data structures
   2. Plotting
   3. Reading/writing files
   4. Conditional logic & control flow
   5. Function definitions
2. Downstream RNA-Seq analysis –
   1. Differential gene expression with DESeq2
   2. Volcano plots & heatmaps
   3. Overrepresentation analysis
   4. Gene Set Enrichment Analysis (GSEA)
   5. PPI-network analysis
   6. Sample level analysis with PCA
3. Upstream RNA-Seq analysis –
   1. Linux shell basics
   2. Seq quality controls with FastQC
   3. Retrieving genomic annotations with UCSC
   4. Alignment and quantification with RSEM
   5. Visualization with IGV

**Course outline**

**Programming with R:**

**Week 1: R, RStudio, and Data structures in R**

Lecturer: Henry Miller

Activity/Homework: Download R & RStudio, Introductory R course in DataCamp

Date: Sept 1st

**Week 2: Conditional logic and control flow**

Lecturer: Paulino Ramirez

Activity/Homework: Intermediate R course in DataCamp, importing flat files chapter in DataCamp, and R assessment in DataCamp, week two practice problems

Date: Sept 8th

**Week 3: Week One and Two Review**

Lecturer: Henry Miller

Activity/Homework: Introduction to the Tidyverse in DataCamp

Date: Sept 15th

**Week 4: The Tidyverse**

Lecturer: Henry Miller

Activity/Homework: Week four practice problems, continue practicing on DataCamp

Date: Sept 22nd

**Week 5: R for biological data science**

Lecturer: Henry Miller

Activity/Homework: Analyze qPCR data using R, RNA-Seq with R course in DataCamp

Date: Sept 29th

**Downstream RNA-Seq analysis with R:**

**Week 6: Differential gene expression**

Instructor: Dr. Yidong Chen

Activity/Homework: Differential gene expression analysis with “airway” data set

Date: Oct 6th

**Week 7: RNA-Seq visualizations**

Lecturer: Dr. Evelien Bunnik

Activity/Homework: EDA and DEG plots exercise

Date: Oct 13th

**Week 8: Biological interpretation**

Lecture: Henry Miller

Activity/Homework: Perform GSEA and ORA analysis exercise

Date: Oct 20th

**Week 9: Databases and Web-based analysis tools**

Lecturer: Henry Miller

Activity/Homework: Gather data from Recount2 and analyze using previous methods along with web-based tools (e.g., enrichr). Complete the Introduction to Shell course in DataCamp.

Date: Oct 27th

**Upstream RNA-Seq analysis with Linux shell:**

**Week 10: Linux shell basics**

Lecturer: Dr. Bernard Fongang

Activity/Homework: Finish the Introduction to Shell course in DataCamp and download fastq files.

Date: Nov 3rd

**Week 11: Alignment and read quantification**

Lecturer: Dr. Chris Chiu

Activity/Homework: Use Salmon to align fastq files and generate count matrices. Import Salmon count matrices with tximport, perform DGE analysis steps.

Date: Nov 10th

**Week 12: Beyond gene expression**

Lecturer: Dr. Siyuan Zheng

Activity/Homework: Start final project.

Date: Nov 17th

**Workshop final project:**

**Week 13: Work on final project**

Activity/Homework: Continue working on projects and preparing presentations

Date: Dec 1st

**Week 14: Presentations and discussion**

Activity/Homework: Student presentations.

Date: Dec 8th

**Additional information:**

* This workshop does not provide a grade and is not credit-bearing.
* **Zoom lectures will be recorded**
* All UTHSA students are welcome. There are no mandatory pre-requisites for this course.
* For additional help, instructors may be reached by email at any time.
* Course materials, including zoom recordings, will be hosted on Box and a link will be provided to all attendees.