# January – May 2025

| | | | |
|---|---|---|---|
| **WEEK 1** | M | 1-13 | First day of classes: <u>Introduction to command line</u><br><br>Students will gain an understanding of the basic skills required to use the command line interface, including how to log into to the Vermont Advanced Computing Core (VACC), a high-performance computing cluster. |
| | W | 1-15 | <u>Introduction to command line continued</u> |
| **WEEK 2** | M | 1-20 | Martin Luther King Holiday |
| | W | 1-22 | <u>Navigating and working with files</u><br><br>In a file system, files are organized into directories or folders which can be further divided into subdirectories. This week students will be introduced to the file system, how to navigate the file system, as well as, manipulating files and directories will be covered. |
| **WEEK 3** | M | 1-27 | <u>Writing, searching, and creating shell scripts</u><br><br>An overview of text editors (`vim, nano`), searching and working with text files, pipe operators, environment variables, and creating shell scripts will be covered. |
| | W | 1-29 | <u>Writing, searching, and creating shell scripts continued</u> |
| **WEEK 4** | M | 2-03 | <u>Considerations when selecting an NGS dataset to analyze from GEO</u><br><br>Experimental planning and considerations when selecting an NGS dataset to analyze will be covered. In addition, how to access a publicly available dataset from NCBI Gene Expression Omnibus (GEO) and job submissions using SLURM, an open-source workload manager designed for LINUX clusters will be described. |

| | | | |
|---|---|---|---|
| | W | 2-05 | **NGS Data Management**<br><br>Overview of best practices with NGS data management and popular tools used for transferring data (Filezilla) will be taught. Description of file formats commonly used in NGS (FASTQ, SAM/BAM) will be provided |
| **WEEK 5** | M | 2-10 | **Preprocessing and Quality Control of sequence reads**<br><br>FASTQC, a popular open-source tool for quality control of high-throughput sequencing data will be covered. |
| | W | 2-12 | **Genome Indexing and Alignment: FASTQ to BAM**<br><br>Introduction to genome indexing and alignment algorithms. Provide an overview of tools (STAR, HiSAT2, BWA-mem) commonly used. Hands-on exercises in genome indexing and alignment will be performed during class. Post-alignment processing to remove low-quality reads, adapter sequences, and contaminants will also be covered. |
| **WEEK 6** | M | 2-17 | Presidents' Day Holiday |
| | W | 2-19 | FASTQ to BAM continued |
| **WEEK 7** | M | 2-24 | **RNA-Seq: BAM to counts files**<br><br>This week we will process BAM files using HTSeq. HTSeq is a python package that can be used to count the number of reads that overlap with features such as genes or exons. |
| | W | 2-26 | BAM to counts continued |
| **WEEK 8** | M | 3-03 | **The R basics: A Review**<br><br>Students have previously been exposed to R programming in the prerequisite course, MMG3310. We will review R syntax (variables, functions, arguments), and data wrangling (data structures, vectors, data frames). |
| | W | 3-05 | **ggplot2**<br><br>Overview of the grammar and how ggplot2 is implemented, the building blocks of ggplot2 plot. |
| **WE** | M | 3-10 | Spring Recess |

| | | | |
|---|---|---|---|
| | W | 3-12 | Spring Recess |
| **WEEK 10** | M | 3-17 | <u>ggplot2 continued</u><br><br>How to create plots including scatterplots, bar charts, and box plots will be covered. Students will come to understand aesthetics, scales, and coordinate systems of `ggplot2`. |
| | W | 3-19 | <u>R and the tidyverse</u><br><br>Learn how to filter, arrange, summarize, mutate and visualize your data with dplyr and ggplot2. |
| **WEEK 11** | M | 3-24 | Dr. Rodriguez away in conference |
| | W | 3-26 | Dr. Rodriguez away in conference |
| **WEEK 12** | M | 3-31 | <u>Replicating Results from Paper #1 in R: Advanced RNA-Seq</u> |
| | W | 4-02 | Paper #1 : RNA-Seq analysis in R continued (HSRF 400) |
| **WEEK 13** | M | 4-07 | <u>Replicating Results from Paper #2 in R: ChIPSeq</u> |
| | W | 4-09 | Paper #2 : ChIPseq analysis in R continued |
| **WEEK 14** | M | 4-14 | <u>Replicating Results from Paper #3 in R: Metagenomics</u> |
| | W | 4-16 | Paper #3 : Metagenomics analysis in R continued |
| **WEEK 15** | M | 4-21 | Catch-Up Week<br><br>We will use this week to "catch-up" on any analysis required for presentation. Students will be required to check-in with Dr. Rodriguez for a 30-min 1-on-1. Sign-up sheet with times will be made available one week prior. |
| | W | 4-23 | Catch-Up Week |
| **WEEK 16** | M | 4-28 | Class Presentations |
| | W | 4-30 | Class Presentations - Last day of classes |