

COMP 383/488 Computational Biology

Quick Start

Spring 2018

Loyola University Chicago

hwheeler1@luc.edu

Bioinformatics Degree Courses



Biology



**Computer
Science**

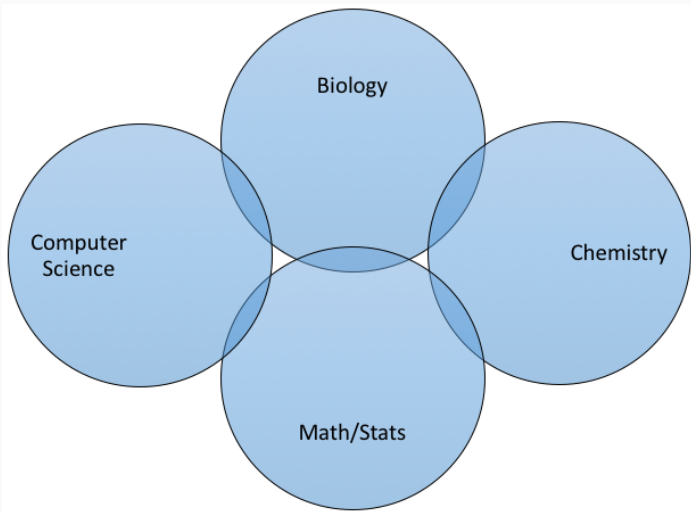


**Math /
Statistics**

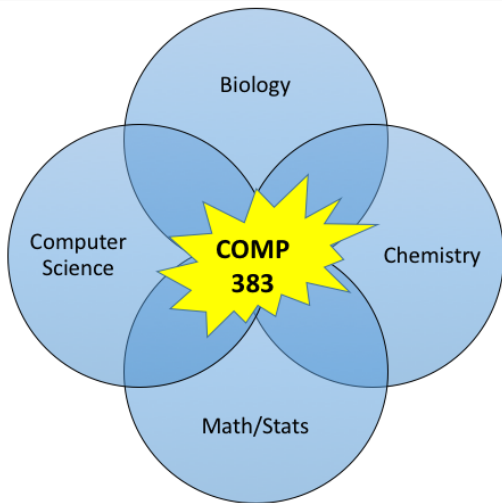


Chemistry

In BIOL 388/488 Bioinformatics, you started to see an intersection of these fields



COMP 383/488 Computational Biology integrates all you've learned



Research must be reproducible

<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003285>

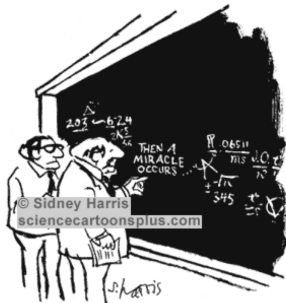
OPEN ACCESS Freely available online



Editorial

Ten Simple Rules for Reproducible Computational Research

Geir Kjetil Sandve^{1,2*}, Anton Nekrutenko³, James Taylor⁴, Eivind Hovig^{1,5,6}



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

Rule 1: For Every Result, Keep Track of How It Was Produced

Rule 2: Avoid Manual Data Manipulation Steps

Rule 3: Archive the Exact Versions of All External Programs Used

Rule 4: Version Control All Custom Scripts

Rule 5: Record All Intermediate Results, When Possible in Standardized Formats

Rule 6: For Analyses That Include Randomness, Note Underlying Random Seeds

Rule 7: Always Store Raw Data behind Plots

Rule 8: Generate Hierarchical Analysis Output, Allowing Layers of Increasing Detail to Be Inspected

Rule 9: Connect Textual Statements to Underlying Results

Rule 10: Provide Public Access to Scripts, Runs, and Results

COMP 383 Course Structure

1. First 5 weeks, learn to apply the Unix command line, Python and R to common biological problems (toy examples) - 30%
2. Group **Research** Project: You will be given freedom to work on a computational biology problem - 70%
 - This could involve automating a research pipeline, building an analysis tool or database, or comparing tools
 - This is a “real life” research project, so data will be messy and unanticipated obstacles will occur
 - It will be important to set benchmarks and stay on task in class
 - Don't procrastinate, get things done and strive to make your code better
 - Learn all you can, the future of biology/medicine/data analysis needs you!

The syllabus and assignments are posted on the class website

<http://hwheeler01.github.io/CompBio/>