

Cbb752b23 Quiz #1 Study Guide

**Quiz #1 will be on Monday Feb 27, 2023 at 1:00pm in Bass 305.
The quiz will be closed-book and last for the whole lecture time.**

There are **some topics** from the lectures that we **drilled down** into detail on. Here is a list of them that you should know for the quiz:

- How to do a dynamic program sequence alignment (i.e. Smith Waterman)
- Principles of alignment and how different methods compare
- How to compute sequence variability (with entropy)
- Proteomics
- How to use simple hash tables & indexes to speed up alignment
- Table normalization + insert/delete/update/select statements
- Definitions of accuracy, sensitivity, specificity, PPV, FPR, and FDR
- PWM and probability matrices
- How to do variant calling SNV, SV
- SVD, PCA, biplots
- Supervised machine learning techniques (i.e. decision trees, SVM)
- Unsupervised machine learning algorithms (i.e. PCA)
- How to do build up a profile matrix (e.g. using a simple EM/psi-blast approach)
- Population scale variation
- Pipeline processing and applications of various assays i.e. ChIP-seq, RNA-seq
- Single cell topics i.e. clustering, technical challenges, workflow, etc.

For **all other topics**, you should understand at a **high level**, i.e., the contents of the lecture. Quiz #1 will cover everything from 1/18 to 2/22.

You can find all slides on the class website: <http://cbb752b23.gersteinlab.org/syllabus>, and all recorded lectures on canvas->media library. Examples of previous quizzes may also be helpful to get an idea of what may be on this year's: <http://cbb752b23.gersteinlab.org/quiz>