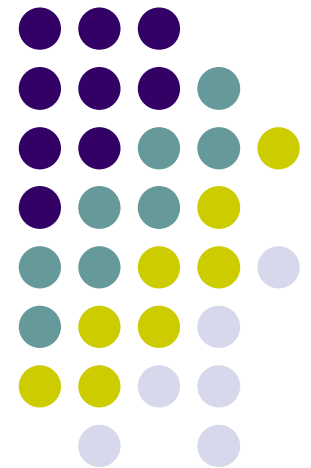


Project Descriptions

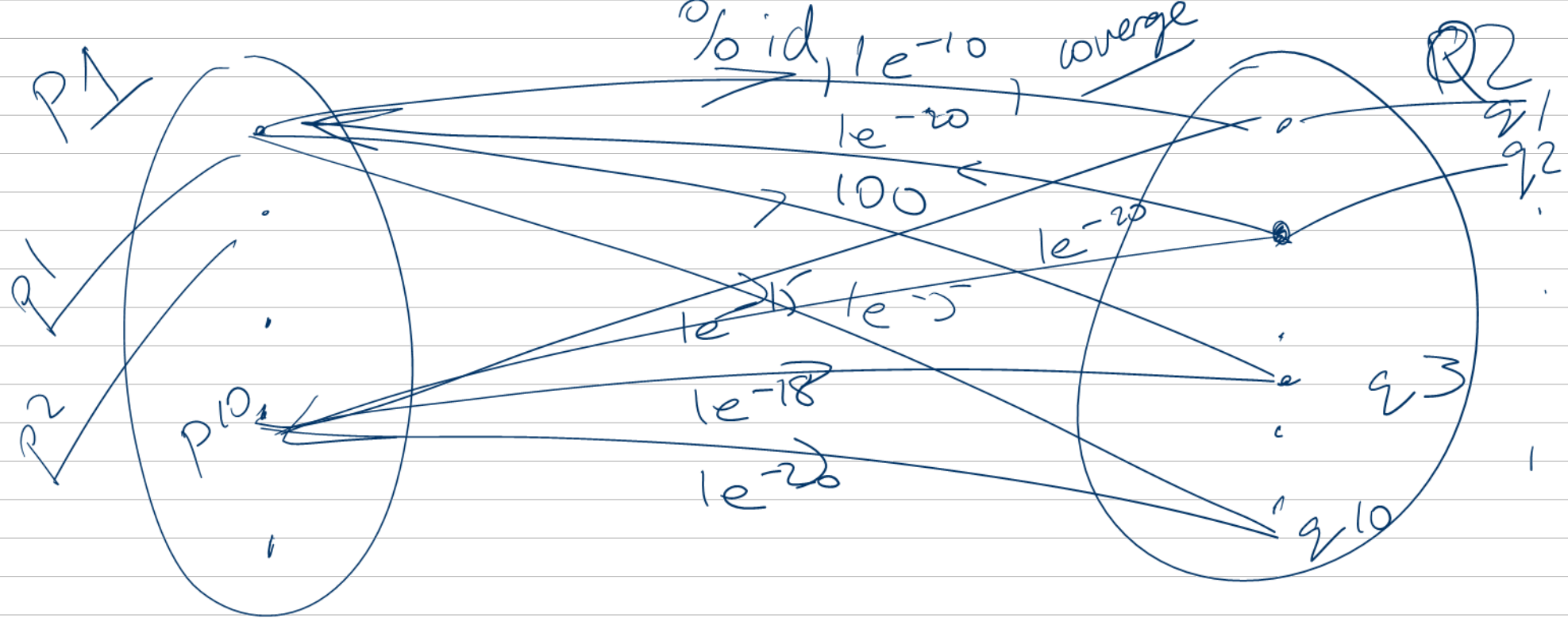
BCHB524
Lecture 20



Class Project: Blast Database

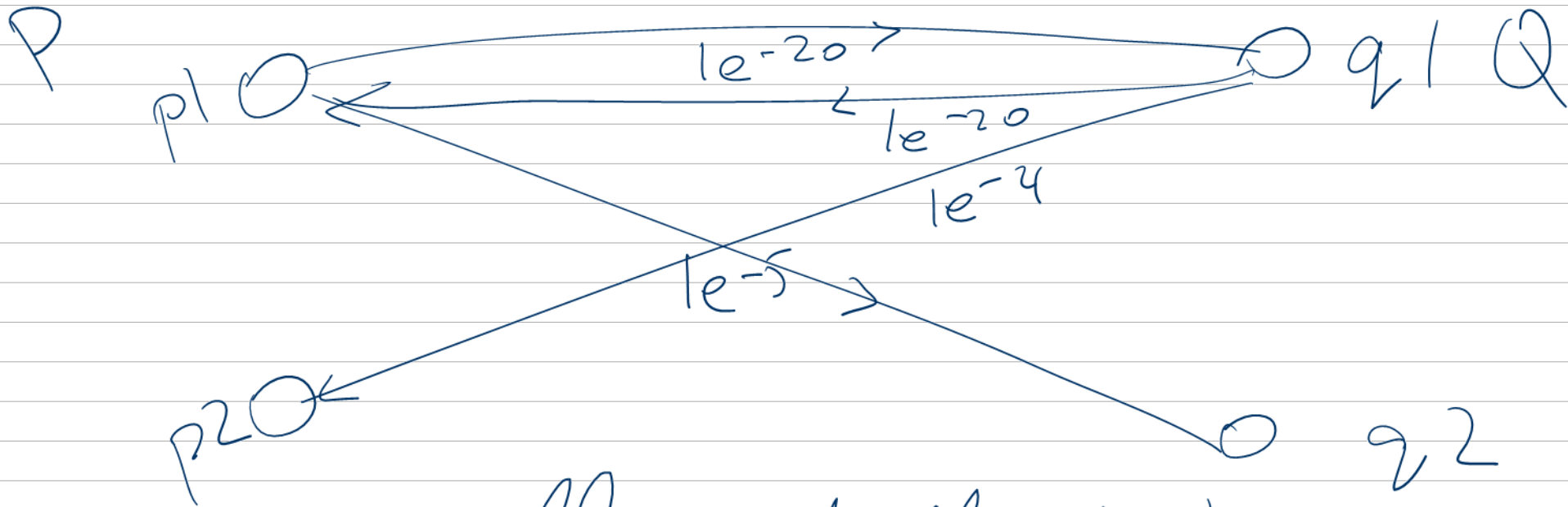


1. Write a program that computes all pairwise blast alignments for two species' proteomes and stores the alignments in a relational database.
2. Write a program that retrieves the blast alignment for two proteins (specified by their accessions) from the relational database.
3. Write a program that finds pairs of orthologous proteins that are mutually best hits in the species' proteomes.

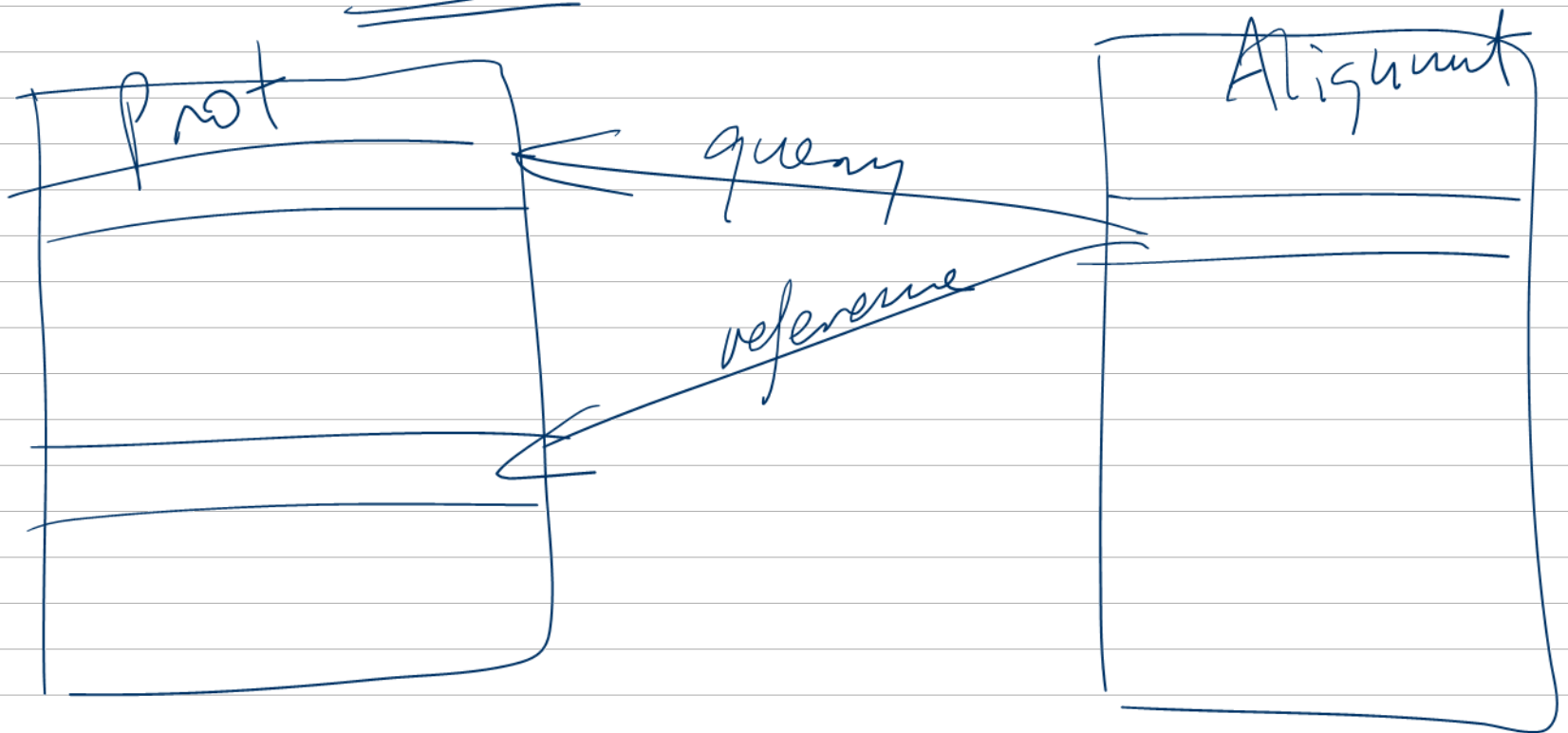


① Small proteomes:
 Yeast ribosome vs ProS ribosome.
 $\leftarrow 5K$ $\leftarrow 80\text{prot}$ $\rightarrow 8K$ $\leftarrow 80\text{prot.}$

② Full small proteomes



all mutually best hits



Class Project: MS/MS Viewer



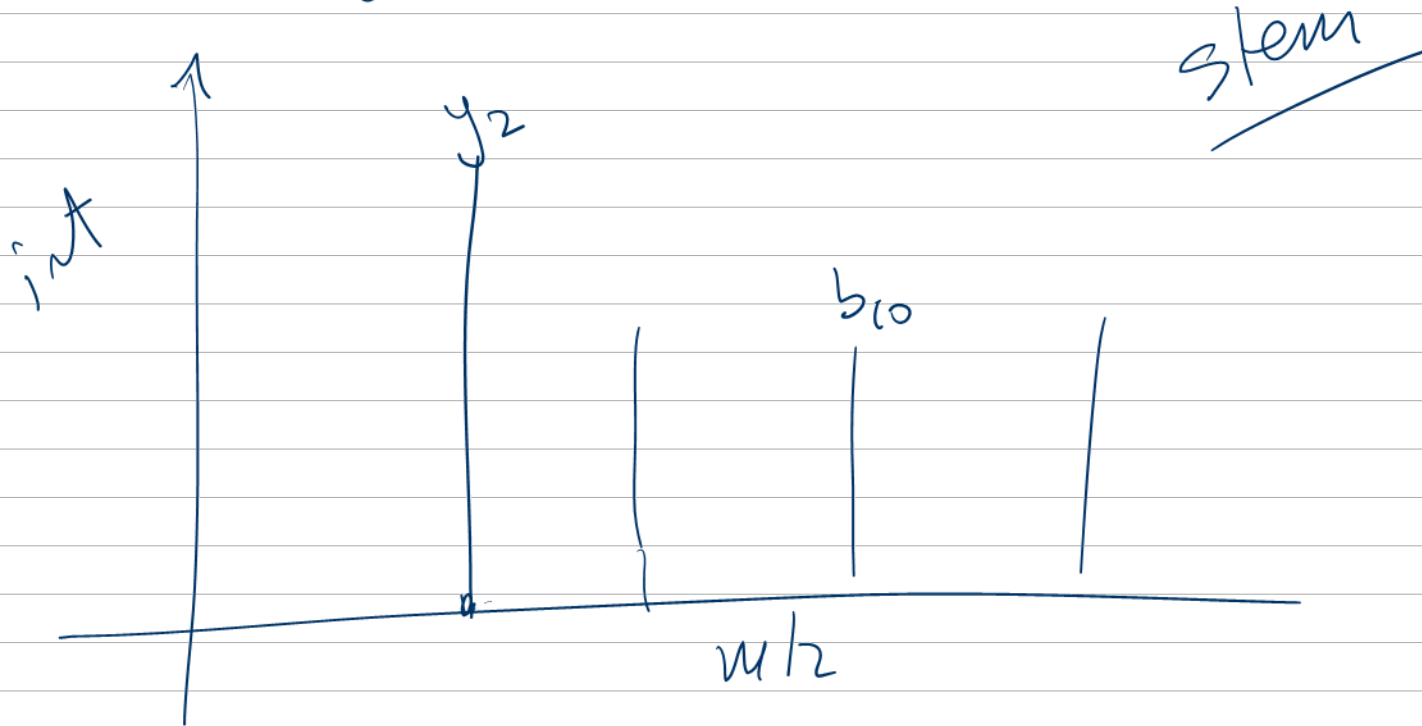
- Write a program to display peptide fragmentation spectra from an mzXML file.
 - The program will take an mzXML file, a scan number, and a peptide sequence as input.
 - The peptide's b-ion and y-ion m/z values should be computed, and peaks matching these m/z values annotated with appropriate labels.
 - The output figure/plot should aid the user in determining whether or not the peptide is a good match to the spectrum.

A C D E F K^l

↑ ←

$$b_i = \sum_{j=1, i}^i mw(aa_i) + \underline{1}$$

$$y_i = \sum_{j=(p-i), l} mw(aa_i) + 19$$





Class Project: Protein Digest

- Write a simple web-server application using TurboGears to carry out an *in silico* enzymatic digest of a user-provided protein sequence.
 - Users should be able to specify min and max length, min and max molecular weight, # of missed cleavages, and specific enzyme.
 - Output should be a table of peptides, with their length, molecular weight, # of missed cleavages, and amino-acids to left and right of each peptide in the protein sequence.

ACDEFK | ACDEF | ACKPEF

$$mw(pq) = \sum_i mw(a_i) + 18$$

- , ACDEFK, A

R. ACDEF, A

R. ACKPEF. -

