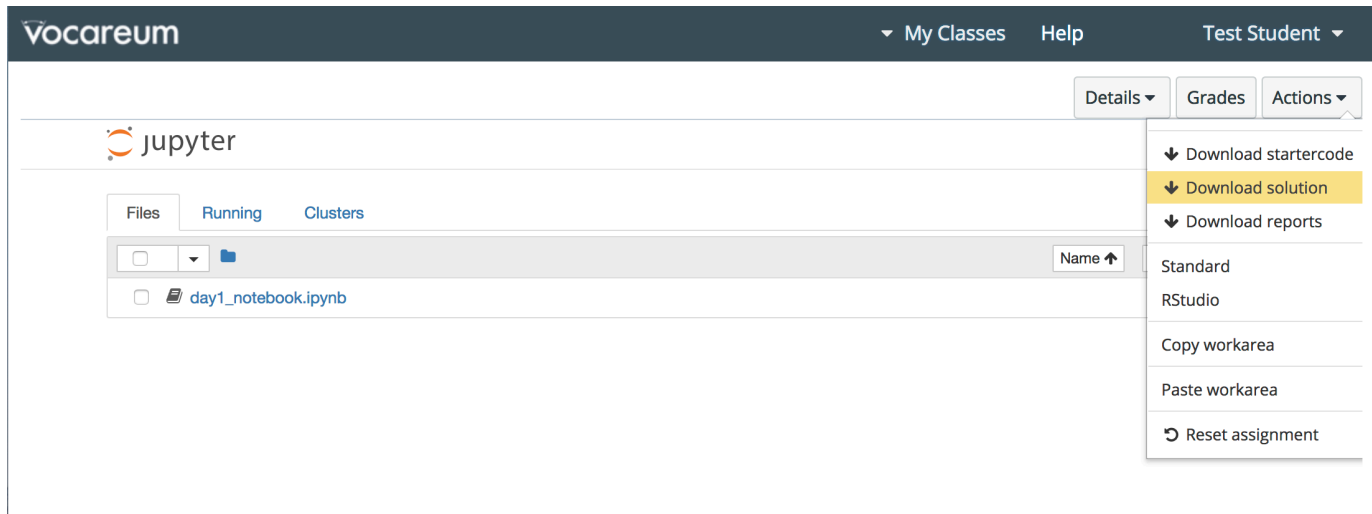


BMI/CS 576 – Day 12

- Today
 - Multiple alignment
 - Scoring
 - Progressive alignment
 - Iterative refinement
- Next week
 - Phylogenetic trees

Reminder: notebook solutions



The screenshot shows the Vocareum web interface. At the top is a dark blue header with the Vocareum logo on the left and navigation links 'My Classes', 'Help', and 'Test Student' on the right. Below the header, there are three tabs: 'Details', 'Grades', and 'Actions'. The 'Details' tab is active. Under 'Details', there is a 'jupyter' logo and a file explorer showing a folder named 'day1_notebook.ipynb'. To the right of the file explorer, there is a dropdown menu with the following options: 'Download startercode', 'Download solution' (highlighted in yellow), 'Download reports', 'Standard', 'RStudio', 'Copy workarea', 'Paste workarea', and 'Reset assignment'.

Options for viewing notebook solutions:

1. Upload into your Vocareum sandbox space
2. Install Jupyter on your personal computer (I recommend the [Anaconda distribution](#))

Consider looking at the solutions even if you got all of the points for a notebook.

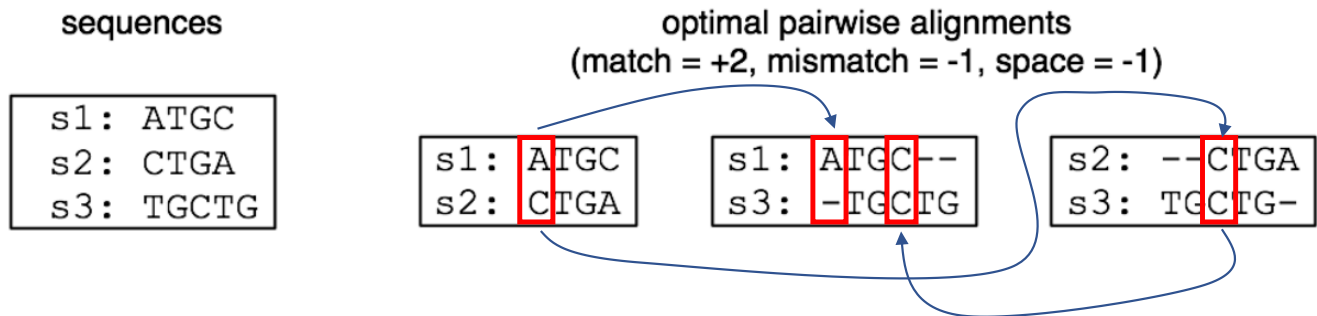
- You might learn a different way to think about a problem or a new Python trick

Midterm

- When: Wednesday, Oct 30th, 5:30-7pm
- Where: 410 Wendt commons (this space)
- What:
 - Sequence Assembly and Sequence Alignment modules
 - Paper exam (no programming)
 - Working with and reasoning about the tasks and their associated algorithms
 - No calculator/electronic devices allowed or needed
 - Two sheets of notes allowed (handwritten recommended!)
 - Old exams available on Canvas

Quiz

Given the three sequences below and all pairwise optimal alignments between them, is there a multiple alignment of the three sequences that is compatible with the optimal pairwise alignments?



No. The aligned pairs are not consistent.

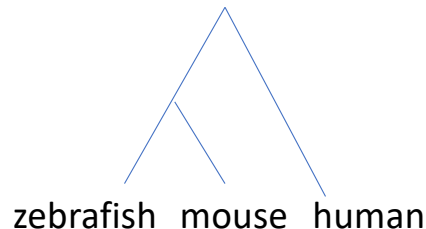
Iterative refinement example

- tiny_seqs in Day 12 notebook

input sequences

zebrafish: SETPKPD
mouse: TEESMYTQI
human: TEGKPTKS

Progressive alignment
with (incorrect) tree:



Alignment (score = 8)

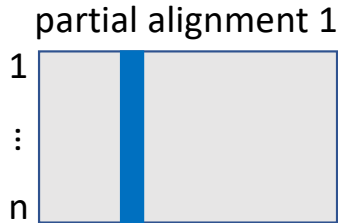
zebrafish: S-E-TPKPD-
mouse: TEE-SMYTQI
human: T-EGKP-TKS

Iterative refinement
(realign first sequence)

Alignment (score = 12)

zebrafish: S-E-TP--KPD
mouse: TEE-SMYTQ-I
human: T-EGKP-TK-S

Scoring an alignment of partial alignments



what is the score of aligning these two columns to each other?



$$S(m_i) = \sum_{k < l} s(m_i^k, m_i^l)$$



$$S(m_i) = \sum_{k < l \leq n} s(m_i^k, m_i^l)$$

Within first alignment



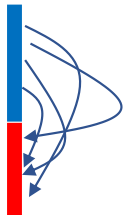
$$+ \sum_{n < k < l \leq N} s(m_i^k, m_i^l)$$

Within second alignment



$$\sum_{k \leq n, n < l \leq N} s(m_i^k, m_i^l)$$

Between two alignments



This is the only term that needs to be computed while aligning the two partial alignments because the other terms are constant with respect to the full alignment