HOW ENCEPHALITIS CAME TO NORTH AMERICA

Tyler Cadigan and Ann Chou

CSC 428: Computational Biology Algorithms

Overview

- Problems
  - Biological
  - Computational
- Approaches
  - Naïve
  - Smart
- Results
  - Individual
  - Composite
- Conclusions

Biological Problems

- Trace West Nile virus outbreak
- Understand methods of transmission
- Control Outbreak

Biological Context

- History
  - Mid to late 20th century, widespread in Asia
  - 1999, first appearance in Western Hemisphere
- Transmission
  - Blood
  - Carried by mosquitoes
- Treatment
  - No human vaccine
  - Limited animal vaccines

Computational Problems

- Edit Distance
  - Input: A collection of strings from a given alphabet
  - Output: The number of add/delete/substitute operations needed to make two strings equivalent
- Hierarchical Clustering
  - Input: A collection of points of data
  - Output: Groupings such that similar elements are together
- Visualization
  - Input: A hierarchical clustering
  - Output: A visual representation of the clustering

Naïve Approaches

- Edit Distance
  - Brute force
- Hierarchical Clustering
  - Trial and error
  - Randomly
- Visualization
  - By hand

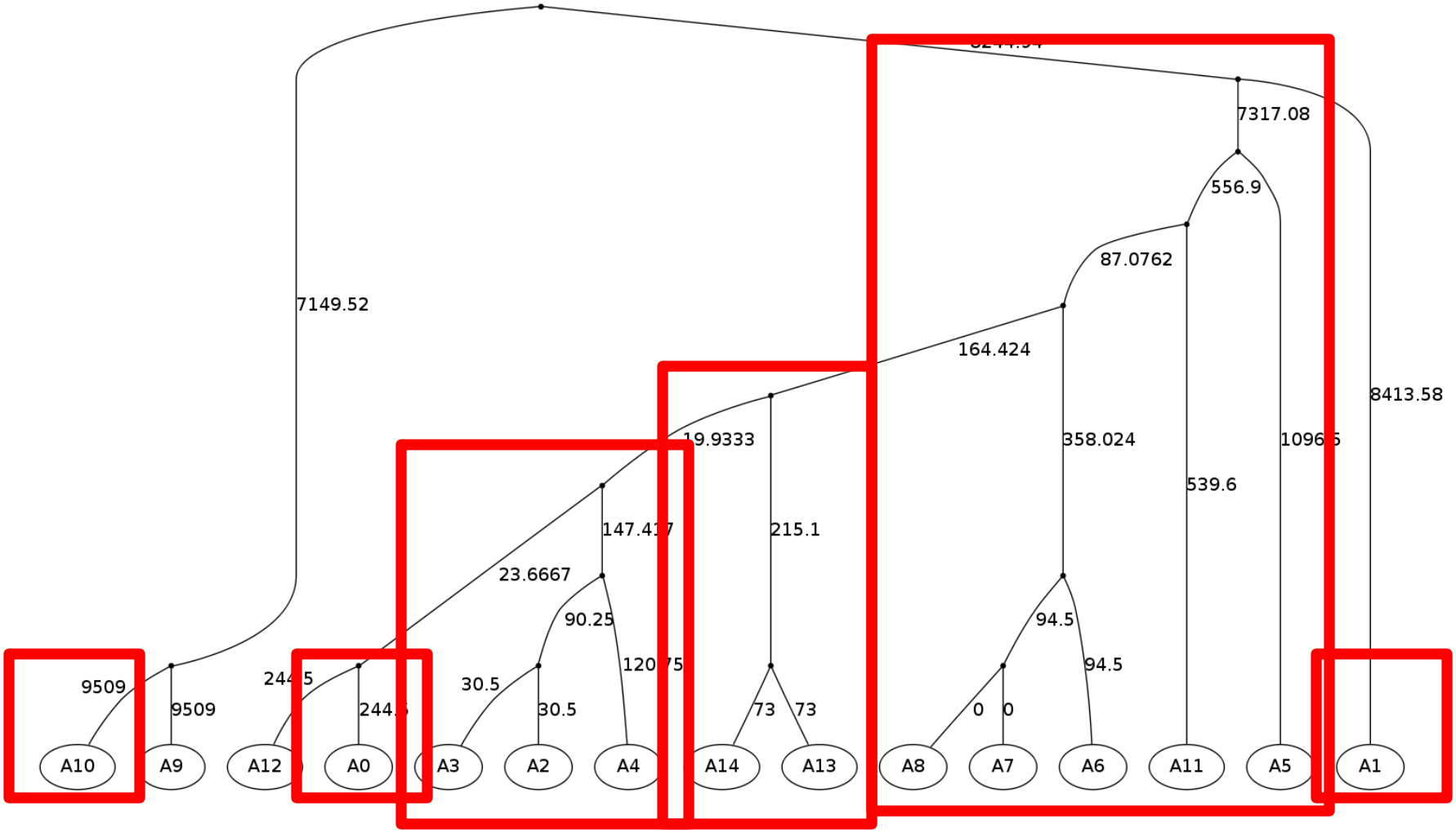
Smart Approaches

- Edit Distance
  - Levenshtein distance
- Hierarchical Clustering
  - UPGMA
- Visualization
  - Dot language

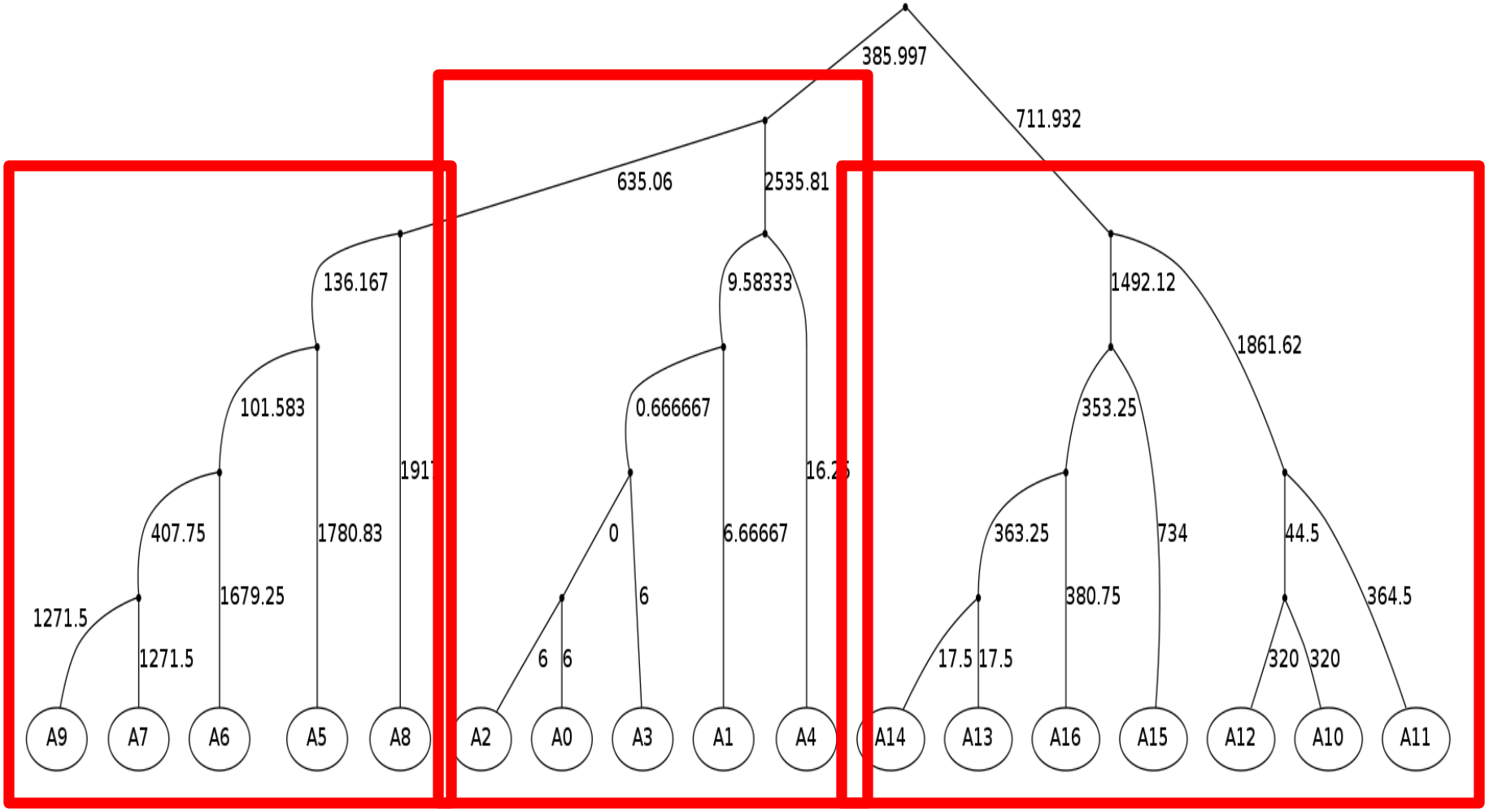
Biological Terms

- Albumin
  - The “servant” in blood
- Mitochondria
  - The “powerhouse” in a cell
- Virus
  - An “enemy” genomic strand

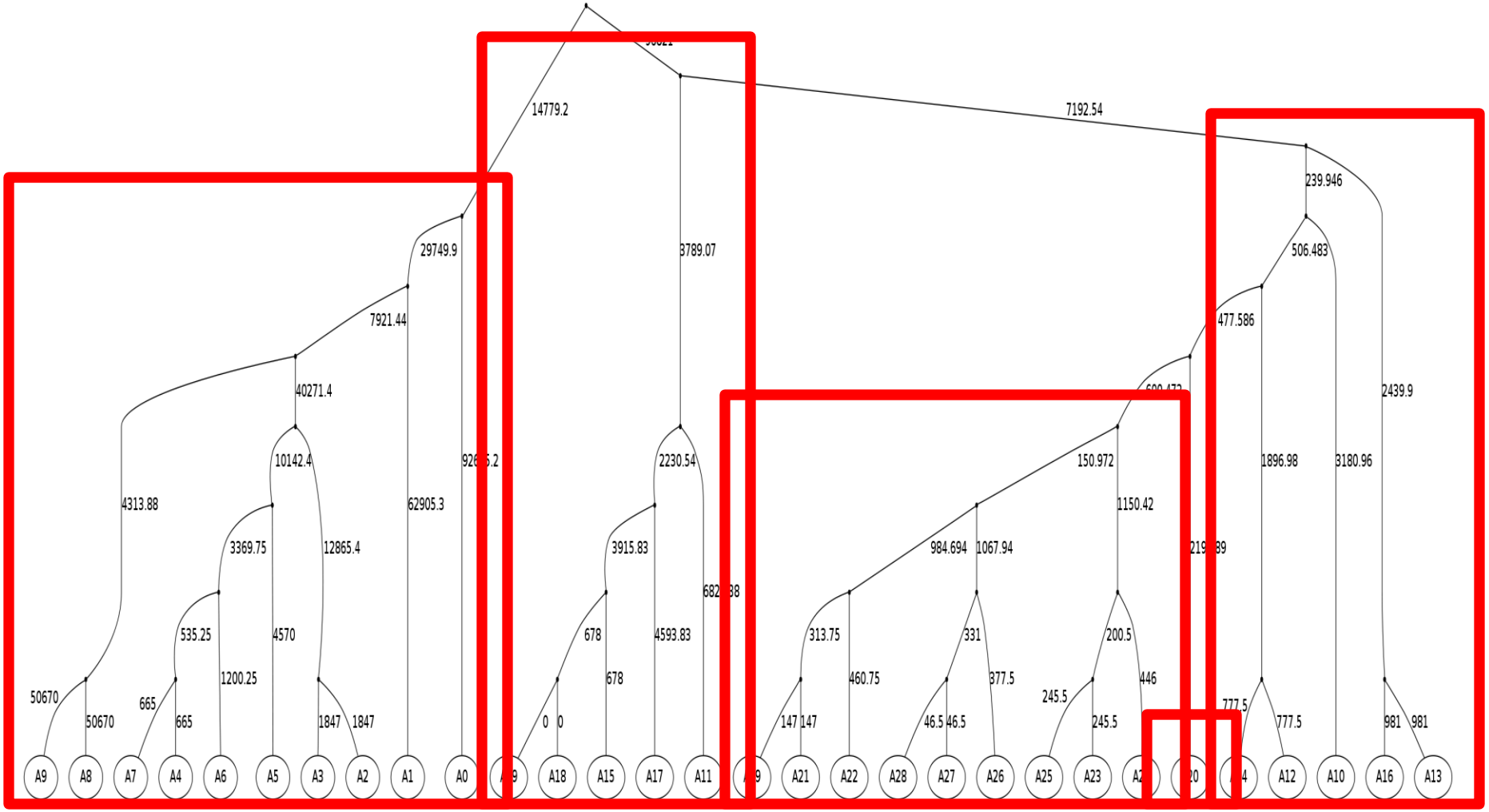
Albumin Results



Mitochondria Results



Virus Results



Complete Topology

Computational Conclusions

- Time consuming
  - ~12 hours for edit distance
- Straightforward
  - Edit distance: 100 lines (C++)
  - UPGMA: 300 lines (Python)
  - Visualization: 200 lines (C++)
  - Supporting programs: 100 lines (C++)

Biological Conclusions

- Mitochondria
  - Humans are closer related to mammals
- Virus
  - Encephalitis is a variant of coronavirus
- Albumin
  - Human albumin has been duplicated several times

Questions?

University of Victoria
Computer Science

Quiz!

- Question 1
  - What were the 3 computational problems faced?
- Question 2
  - What were the 3 biological sequence origins investigated?
- Question 3
  - What type of virus is encephalitis closely related to?

    A) Pox B) Dengue