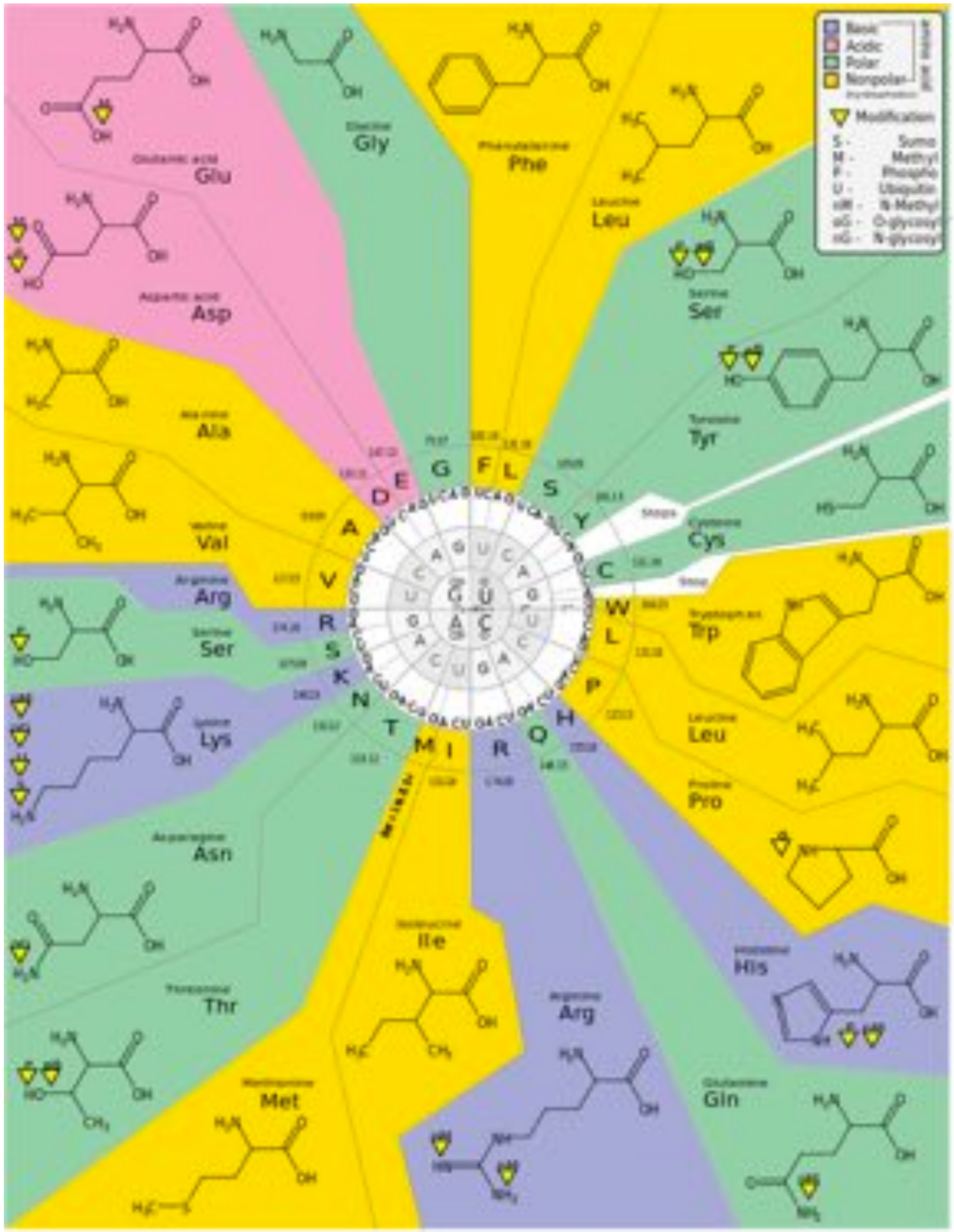
Lecture 17. Protein Structure Prediction

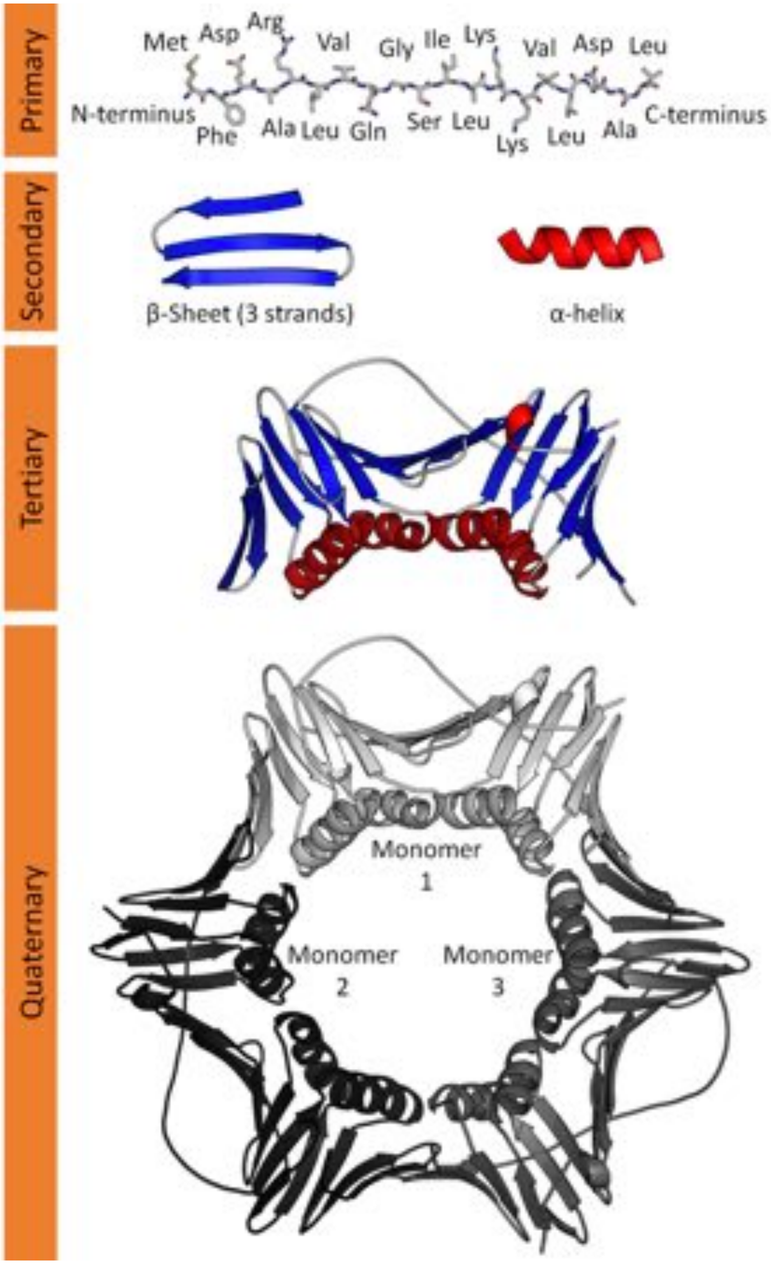
Michael Schatz

March 30, 2020

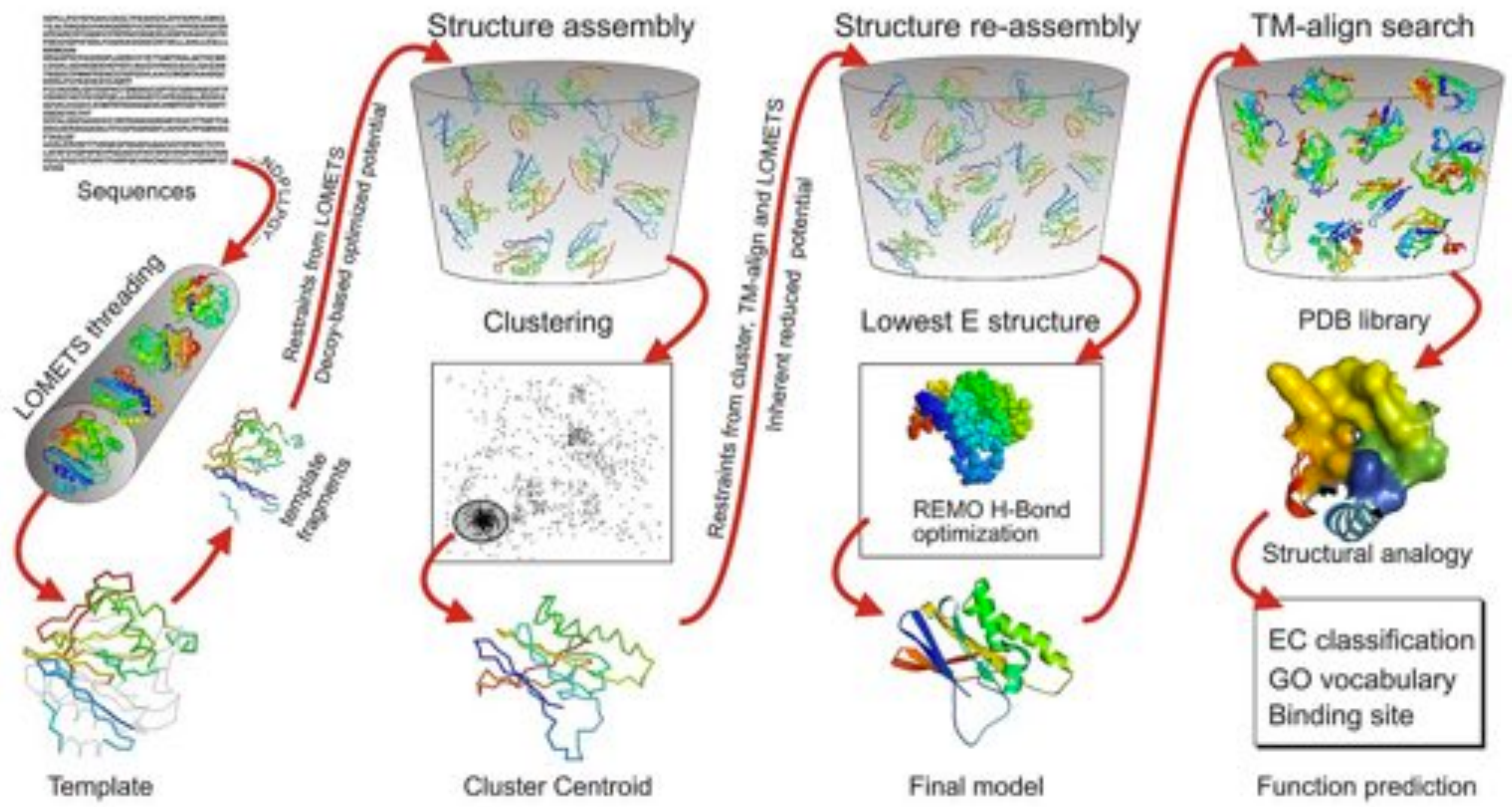
JHU 600.749: Applied Comparative Genomics

Protein Structure



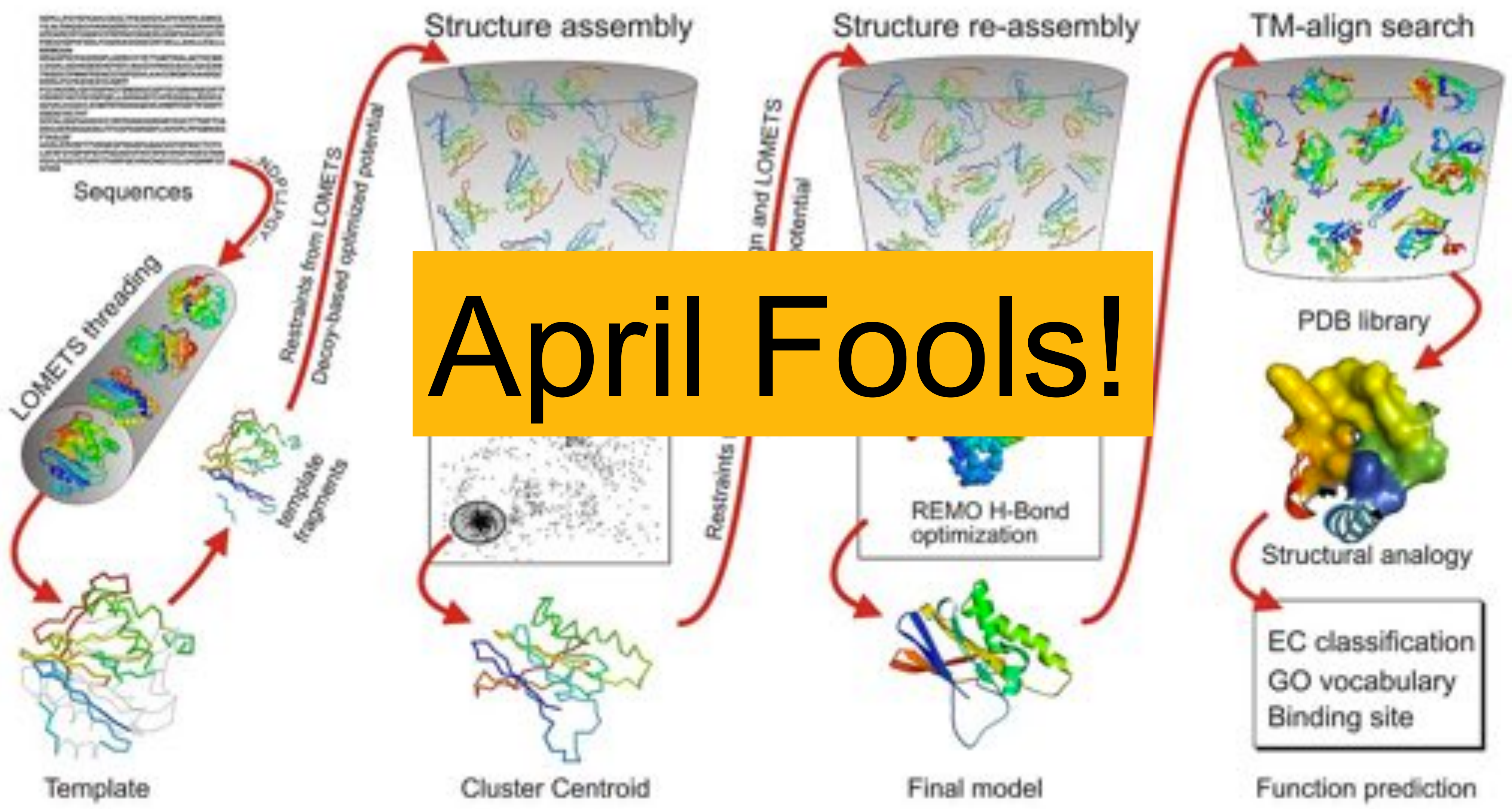


Protein Structure Prediction



I-TASSER: a unified platform for automated protein structure and function prediction
Roy et al (2010) *Nature Protocols* 5, 725–738. doi:10.1038/nprot.2010.5

Protein Structure Prediction



April Fools!

I-TASSER: a unified platform for automated protein structure and function prediction
Roy et al (2010) *Nature Protocols* 5, 725–738. doi:10.1038/nprot.2010.5

Review!

Topics

Genomics

- Genomics Technologies
  - Illumina, PacBio, Nanopore
- Genome Assembly
- Whole Genome Alignment
- Read mapping
- Variant Identification: SNVs - SVs
- Gene Finding
- RNA-seq
- Methyl-seq, Chip-Seq, Hi-C
- Genome Annotation
- ENCODE
- Single cell vs bulk sequencing

Quantitative Techniques

- Normal, Poisson, Binomial, P-value
- de Bruijn and overlap graphs
- Minimizers
- Dot plots
- Quality Values (Phred Scale)
- Full text indexing & BWT
- Seed & Extend
- Hidden Markov Models
- PCA / t-SNE / UMAP
- Differential Expression
- Sampling variance
- Expectation Maximization

What is the goal? What is the approach? What are the key challenges?

How did we explore these topics in the homeworks and lectures?