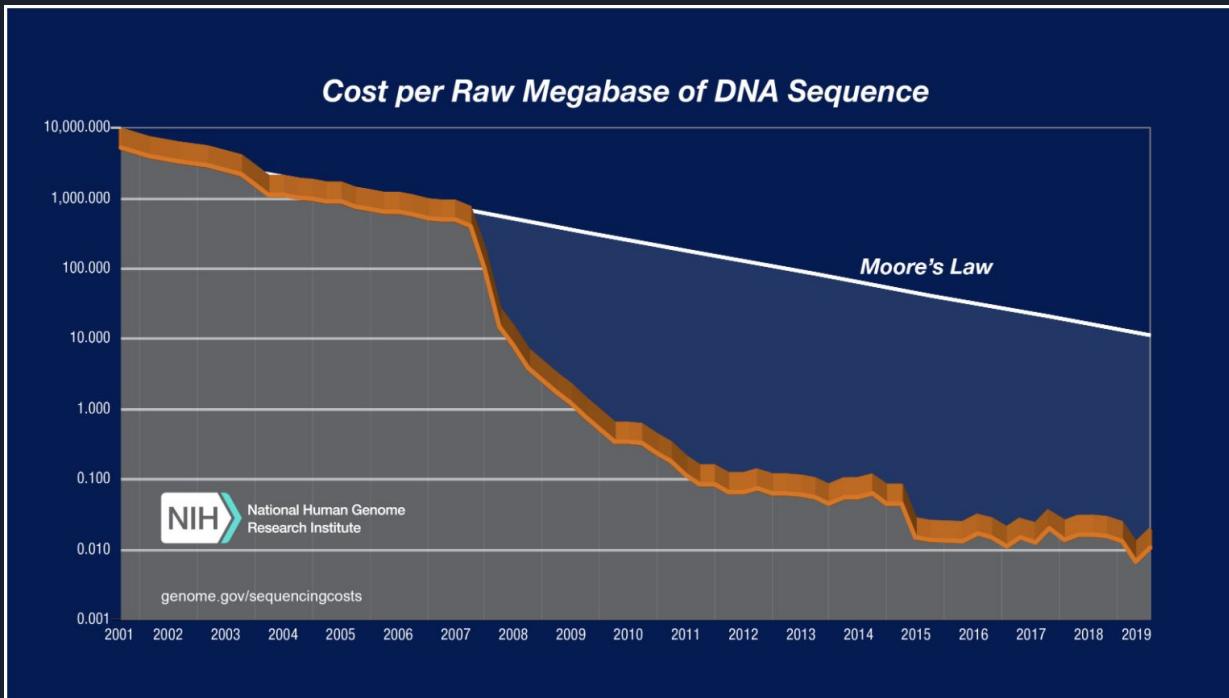




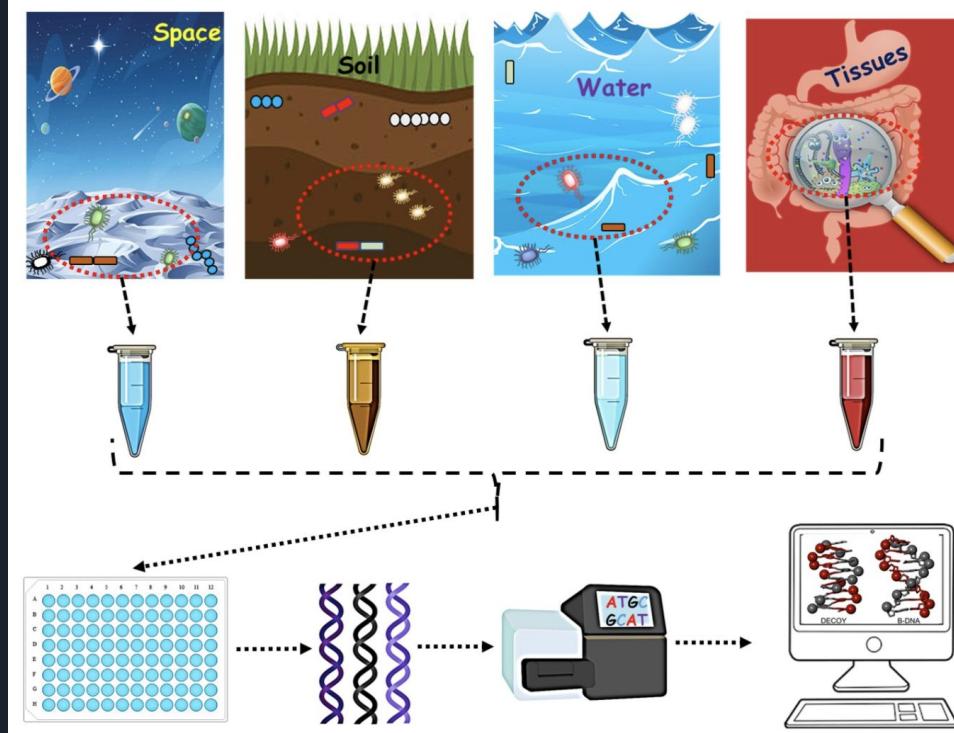
Optimizing Data Processing of Metagenomes

Sergio E. Mares - Oklahoma State University
Humberto G. Ortiz-Zuazaga - University of Puerto Rico, Rio Piedras

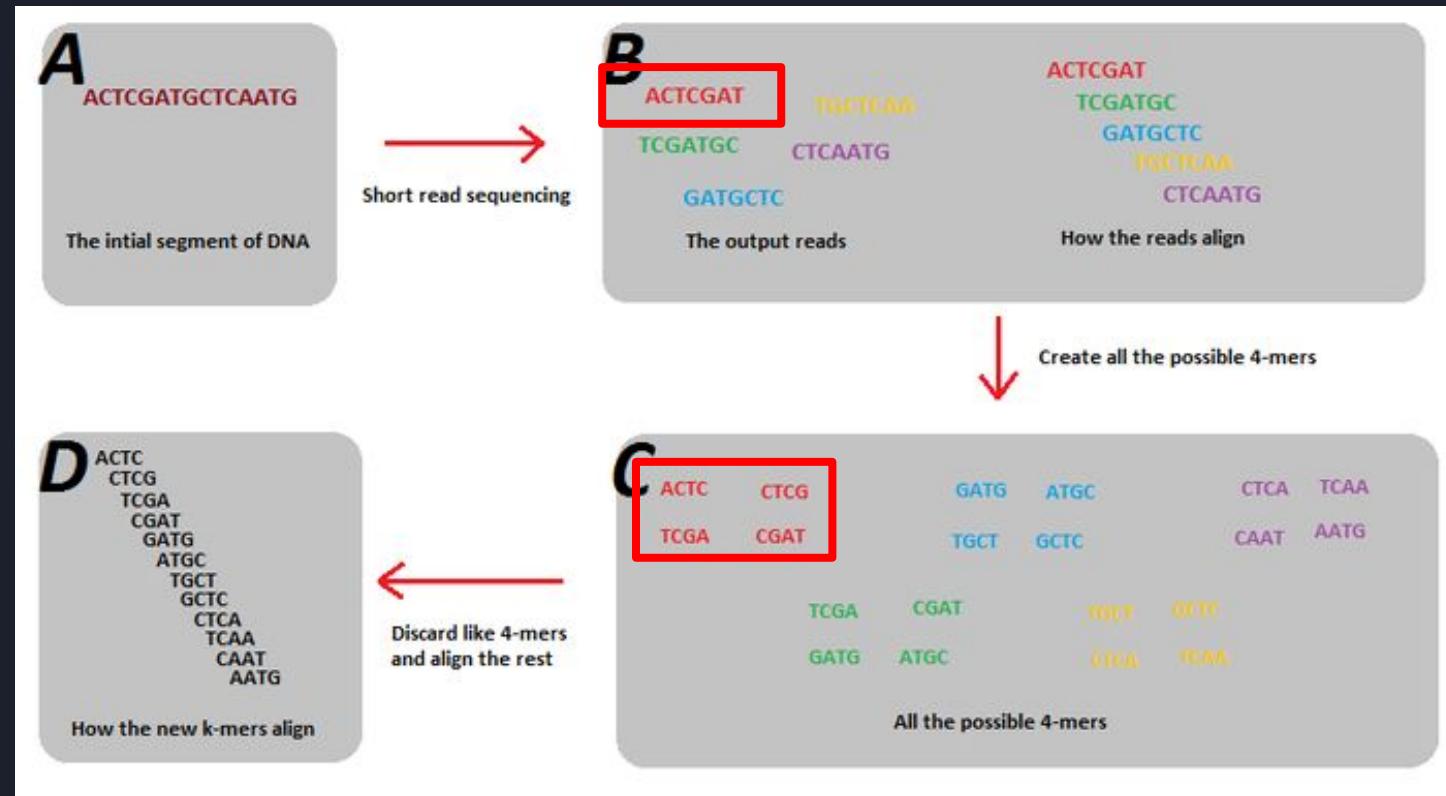
Data cannot be processed quick enough



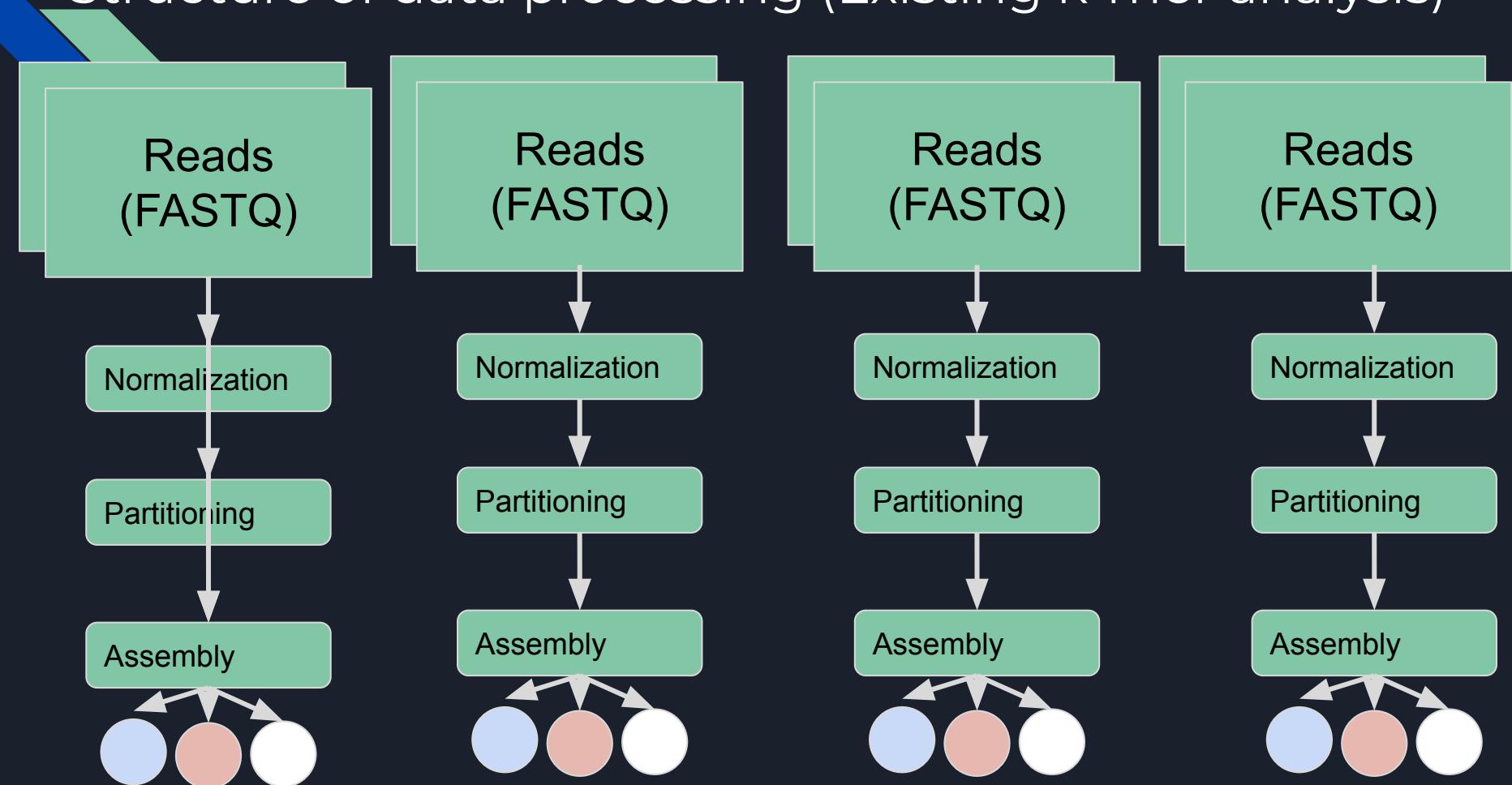
Metagenomics



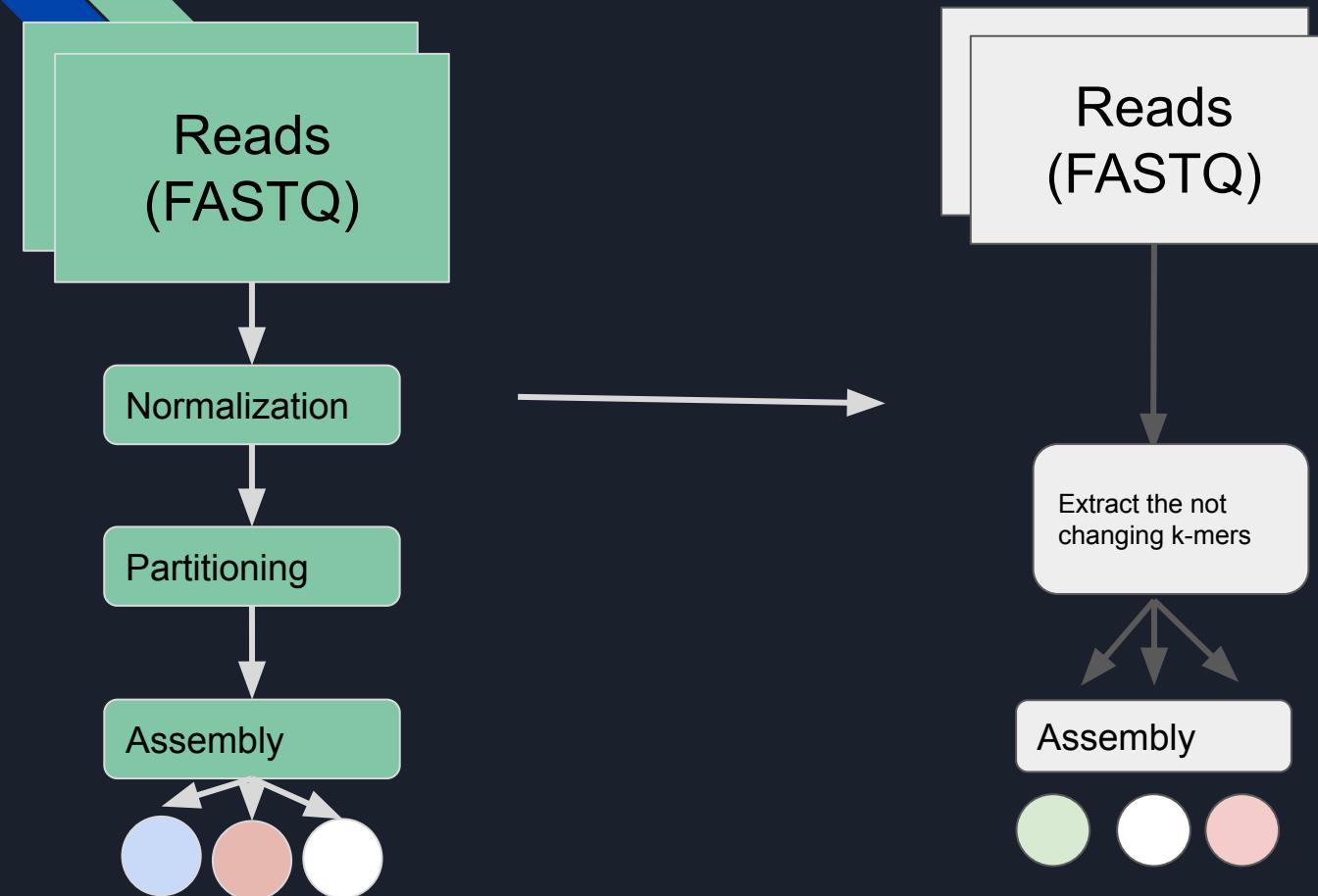
Structure of data processing (Existing k-mer analysis)



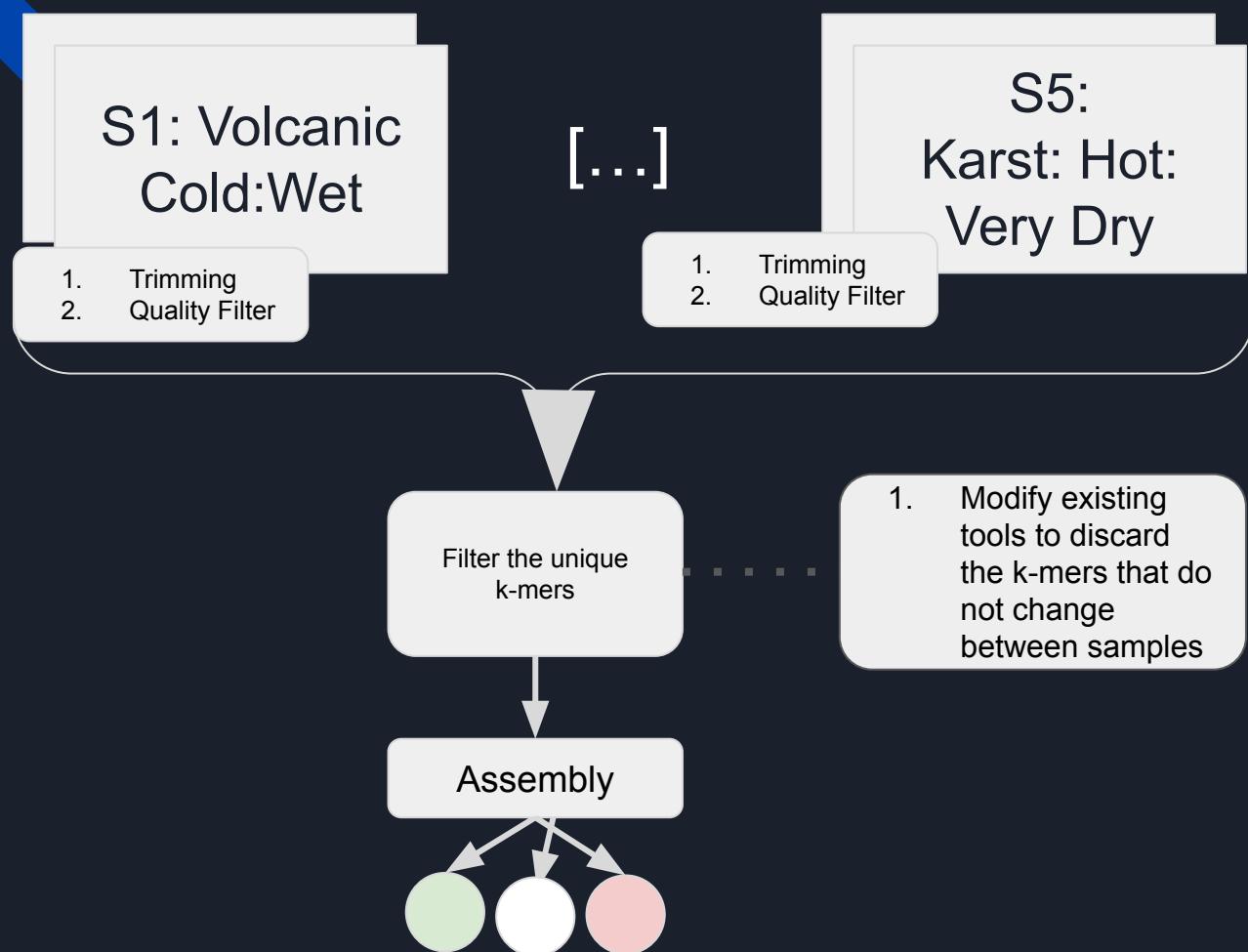
Structure of data processing (Existing k-mer analysis)



A Statistical Approach: ANOVA

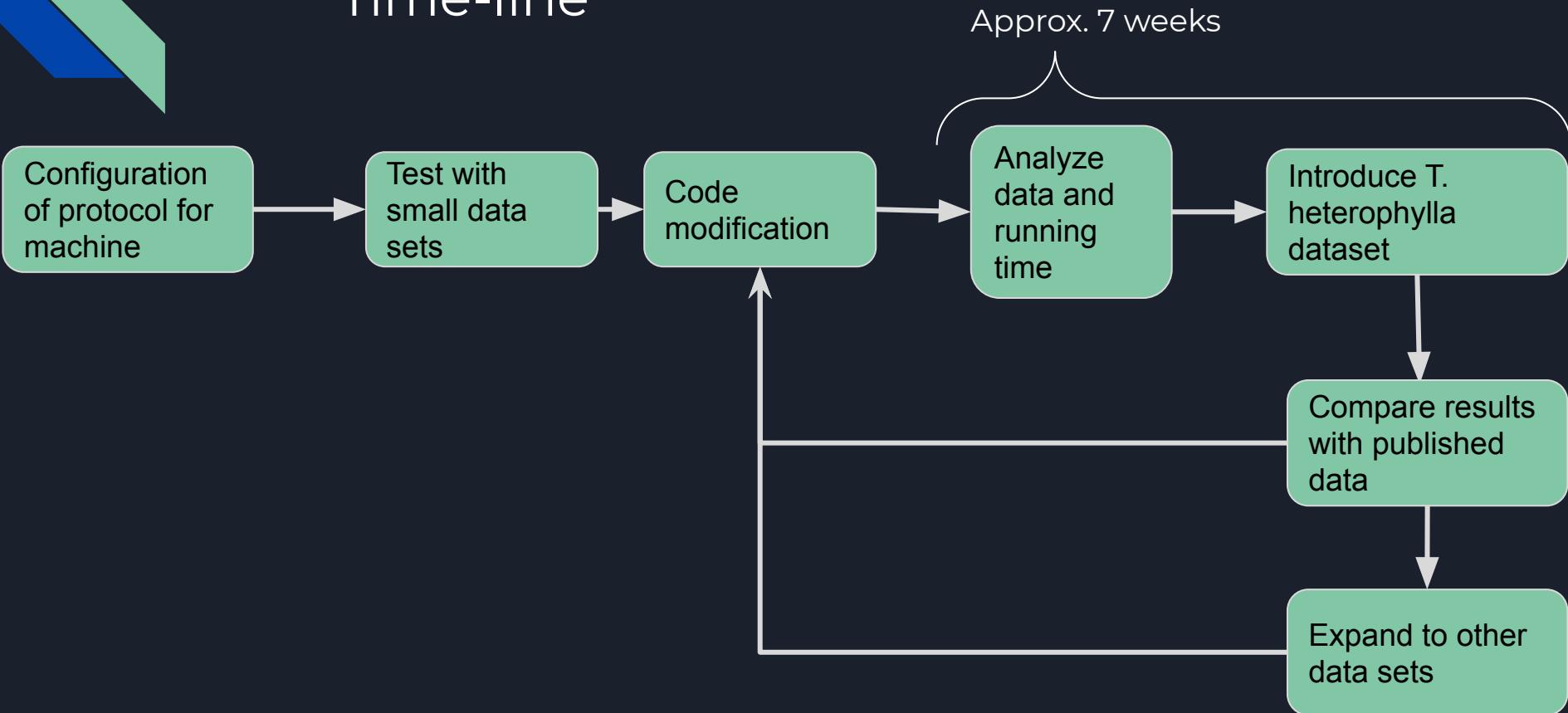


A Statistical Approach: ANOVA



Tabebuia heterophylla

Time-line





Expected Results

	Size of the file (bytes)
Trim (Step 1)	662,643,774
Normalization (Step 2)	148,972,458
Statistical ANOVA Approach	???

Improving data processing for quicker analysis

