

## BIO 310 HW3

Alihan Kerestecioğlu

ID: 25497

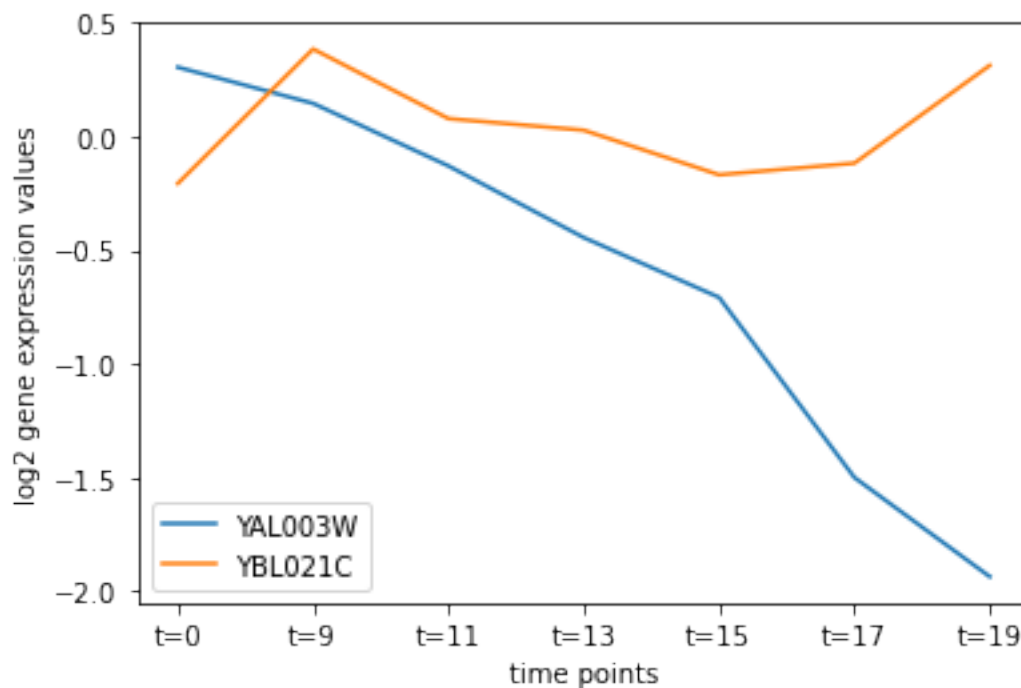
### Generate string compositions and de Bruijn graphs

3. Sequence that the graphs are generated from is the same for all three of them but the  $k$  value varies. As  $k$  increases, the graph is guaranteed to be more unique but the size of the graph (the number of vertices) increases.

4. Although the sequences are different, the graphs generated from these sequences are the same.

### Study of Yeast Gene Express

1.



The log2 gene expression value of the gene with the id YAL003W(blue line) decreases over time while the other gene has a line with up and downs.

2.

The correlation coefficient is -0.14285714285714288. Although it is a negative number, since the coefficient is very close to 0, it can be said that there is almost no correlation.

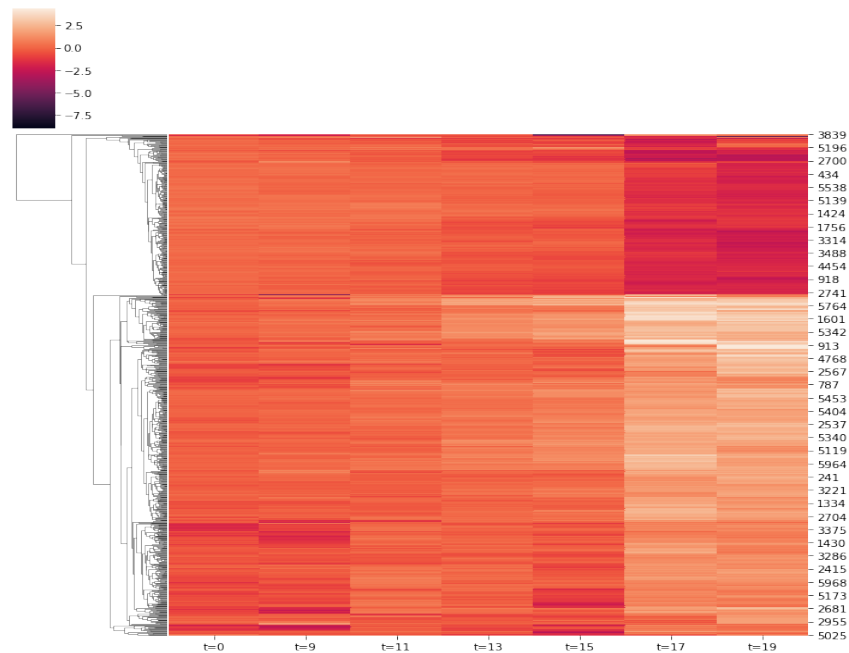
3.

The gene with the id YJL009W varies the most while the gene with the id YPL016W varies the least.

4.

The lowest variance is 0.535379 and it belongs to the gene with the id YDL051W.

5.



I clustered the genes with the average linkage and Euclidian distance function. The genes are clustered well. After t=17 colors vary more.