

## Instructions

1. Find putative homologs of query amino acid sequences in FASTA format (same as last week's instruction - i.e., what your current script does)
2. Use Bash loops and Bash conditional statements (no Python!) to identify which genes in a BED file contain the identified homologous histidine kinase domains. (Hint: if a gene contains a domain, then the location of the domain will be entirely within the boundaries of the gene)
3. Write an output file containing the unique gene names which your script identified as containing predicted HK domains.

Usage of your script should be: `homolog_identify.sh <query.faa> <subject.fna> <bedfile> <outfile>`

```
(HW2) sarthdiskalkar@Sarths-MBP week4_data % ./sdiskalkar3.sh HK_domain.faa Escherichia_coli_K12.fna Escherichia_coli_K12.bed Escherichia_coli_K12_results
matches found in Escherichia_coli_K12.fna
(HW2) sarthdiskalkar@Sarths-MBP week4_data % wc -l Escherichia_coli_K12_results
wc: illegal option -- 1
usage: wc [-clmw] [file ...]
(HW2) sarthdiskalkar@Sarths-MBP week4_data % wc -l Escherichia_coli_K12_results
103 Escherichia_coli_K12_results
(HW2) sarthdiskalkar@Sarths-MBP week4_data % ./sdiskalkar3.sh HK_domain.faa Wolbachia.fna Wolbachia.bed Wolbachia_results
matches found in Wolbachia.fna
(HW2) sarthdiskalkar@Sarths-MBP week4_data % wc -l Wolbachia_results
4 Wolbachia_results

(HW2) sarthdiskalkar@Sarths-MBP week4_data % ./sdiskalkar3.sh HK_domain.faa Vibrio_cholerae_N16961.fna Vibrio_cholerae_N16961.bed Vibrio_cholerae_N16961_results
matches found in Vibrio_cholerae_N16961.fna
(HW2) sarthdiskalkar@Sarths-MBP week4_data % wc -l Vibrio_cholerae_N16961_results
107 Vibrio_cholerae_N16961_results
(HW2) sarthdiskalkar@Sarths-MBP week4_data %

107 Vibrio_cholerae_N16961_results
(HW2) sarthdiskalkar@Sarths-MBP week4_data % ./sdiskalkar3.sh HK_domain.faa Pseudomonas_aeruginosa_UCBPP-PA14.fna Pseudomonas_aeruginosa_UCBPP-PA14.bed Pseudomonas_aeruginosa_UCBPP-PA14_results
matches found in Pseudomonas_aeruginosa_UCBPP-PA14.fna
(HW2) sarthdiskalkar@Sarths-MBP week4_data % wc -l Pseudomonas_aeruginosa_UCBPP-PA14_results
217 Pseudomonas_aeruginosa_UCBPP-PA14_results
(HW2) sarthdiskalkar@Sarths-MBP week4_data %
```

Script is attached in canvas.