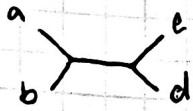
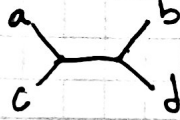


PART 1

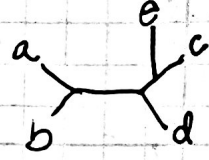
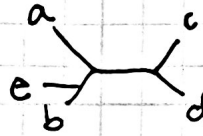
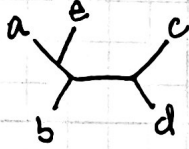
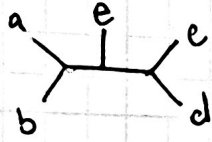
1) 5 taxa: abcde



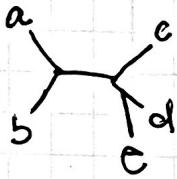
no 4 taxa or



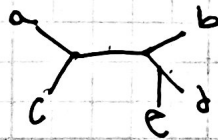
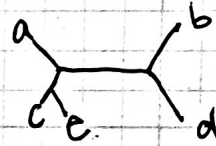
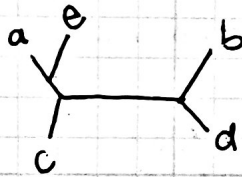
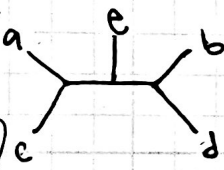
or



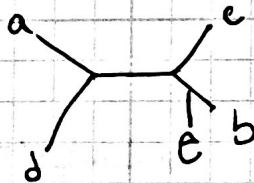
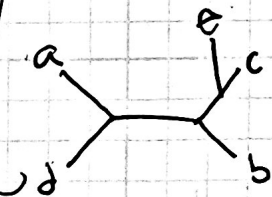
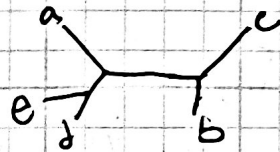
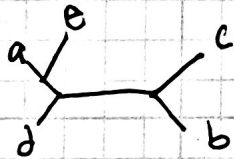
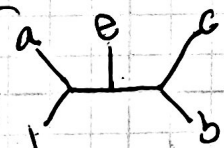
group 1



group 2



group 3



PART 1 cont.

2.

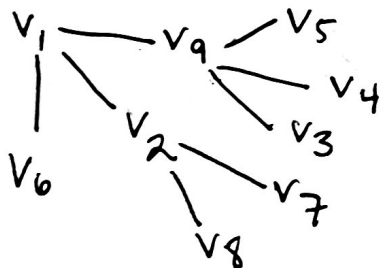


- a.
- $v_1 = 3$ degrees
 - $v_2 = 3$ degrees
 - $v_3 = 1$ degree
 - $v_4 = 1$ degree
 - $v_5 = 1$ degree
 - $v_6 = 1$ degree
 - $v_7 = 1$ degree
 - $v_8 = 1$ degree
 - $v_9 = 4$ degrees

b. $v_4, v_5, v_6, v_7, v_8, v_3$

c. no. There are several unresolved nodes.

d.



e.

$(v_5, v_4, v_3)(v_8, v_7)(v_6))$

github PART 2

<http://github.com/demolnau/EE08563-lab2019.git>

PART 3

a. cloning vector pAgal6 Genbank MH621333.1

b. ~~gallus gallus GATA binding protein~~
~~if using nucleotide~~
 → blastX results → erythroid transcription factor
 (gallus gallus)
 NP_990795.1

The sequences do not match exactly. There are gaps in the sequence of the query and some ~~subsequences~~ that spell MARK WAS HERE WITH ~~the same sequence~~ ~~and the first question~~

There are lighter-green letters that appear that are low complexity sequences that are typically masked.

5. The 20 most similar hits to the dino sequence are homologous to one another. They are orthologs if they are specifically sequences from other species and paralogs if the two copied are similar and within the same species (due to a duplication event).

the erythroid transcription factor *gallus gallus* is ortholog to the erythroid transcription factor of *Lonchura striata domestica*, *python bivittatus*, *panus major*, *pseudobutes humilis*, *Alligator mississippiensis* etc.
The erythroid transcription factor isoform X2 & X1 of *Panus major* are paralog to one another.

PART 4.

~~mafft~~ ~~global pair~~ ~~maxiterate 1000~~ ~~erythroid~~ ~~...~~ ~~.txt~~
~~mafft~~ ~~global pair~~ ~~maxiterate 1000~~ ~~erythroid~~ ~~...~~ ~~.txt~~
~~mafft~~ ~~global pair~~ ~~maxiterate 1000~~ ~~erythroid~~ ~~...~~ ~~.txt~~
mafft
-- global pair
-- maxiterate 1000
-- retree 0
-- treeout
erythroidtxt > eryth ... tree

move from nrc to local
scp demoinan@nrc ... /eryth ... tree
put into figtree to visualize

I used a global build because all sequences were about the same length.

This tree represents the relationship between sequences and attempts to show the relationship as distance. Note that it is unrooted.

