

## CS 175 Problem set 2

Rianna Patricia Cruz (2013-59075)

Relevant files (unprocessed and processed files, and assets) may also be found at <https://github.com/rpscruz/CS175PS/tree/master/CS175PS2>.

1.

Create a table as indicated below, filling it up with the relevant information obtained using the BLAST search. Sequence GenBank Taxon with Closest Taxon Accession No. of Closest Match % Identity E-value

Seq	GenBank Taxon with Closest Taxon	Accession No. of Closest Match	% Ident	E-val
B1	Sus scrofa breed Duroc mitochondrion, complete genome	MF183225.1	99%	0.0
B2	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
B3	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
B5	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
U1	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
U2	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
U3	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0
U5	Sus scrofa isolate yorkshire 1-38-177 breed Yorkshire mitochondrion, complete genome	JN601074.1	99%	0.0

2.

Based on the obtained results, were the meats properly labelled?

No the meats were not properly labelled, as scrofa isolate yorkshire describes the genome of a wild boar, while Sus scrofa breed Duroc describes the genome of a domestic pig<sup>1</sup>.

3.

For this item, place the screenshots of the trees generated, grouping them according to the file used to generate the tree. Is there a concurrence with the trees generated by all of the techniques used for samples B1, B2, B3, and B5? How about for the trees generated using samples U1, U2, U3, and U5?

For this item, FigTree was used to generate 5 phylogenetic trees each for both the U test dataset and the B test dataset, which may be found on the following pages.

For the 'baboy-ramo' dataset.

As it appears that B2, B3, and B5 appear to alternately share a phylogenetic branch in several (n=5) runs, it may be inferred that they are closely related to a significantly similar degree. The 'best tree' generated from the Figtree application also presents these as branching off of the same branch from the phylogenetic tree, while B1 constantly remains on a distinct branch.

This also appears to be the case for U1, U3, and U5 appear to branch off the first level of the Figtree-generated phylogenetic tree, while U2 is alternately closely related (ie. on the same branch) as U3 and U5. It may thus be inferred that the samples may not be so genetically distinct from each other.

For the 'usa' dataset.

As it appears that B2, B3, and B5 appear to alternately share a phylogenetic branch in several (n=5) runs, it may be inferred that they are closely related to a significantly similar degree. The 'best tree' generated from the Figtree application also presents these as branching off of the same branch from the phylogenetic tree, while B1 constantly remains on a distinct branch.

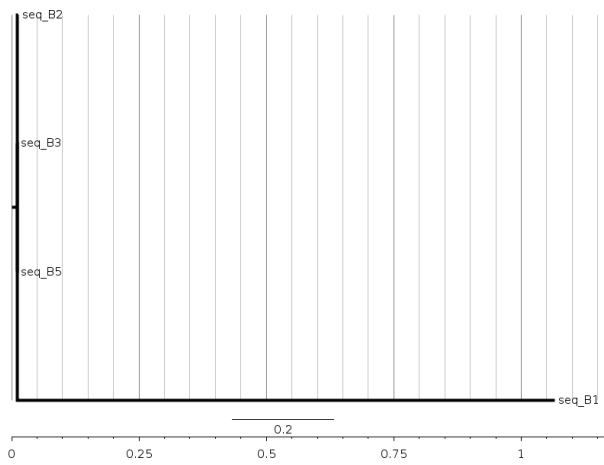
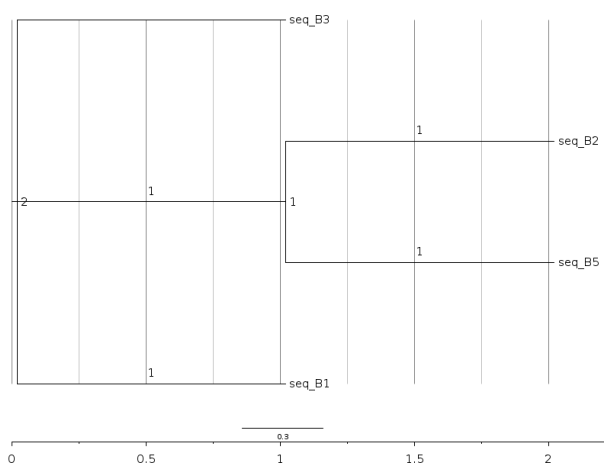
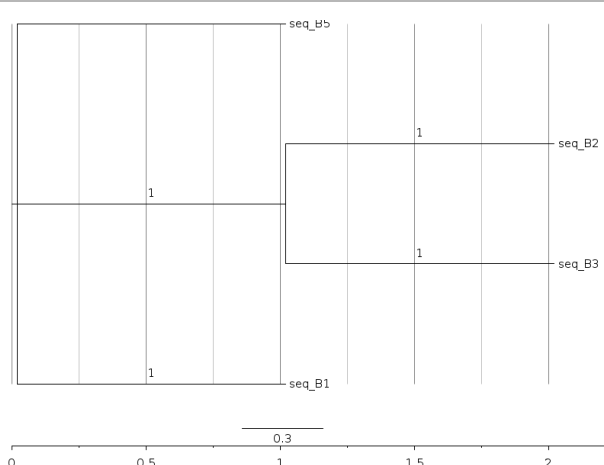
This also appears to be the case for U1, U3, and U5 appear to branch off the first level of the Figtree-generated phylogenetic tree, while U2 is alternately closely related (ie. on the same branch) as U3 and U5. It may thus be inferred that the samples may not be so genetically distinct from each other.

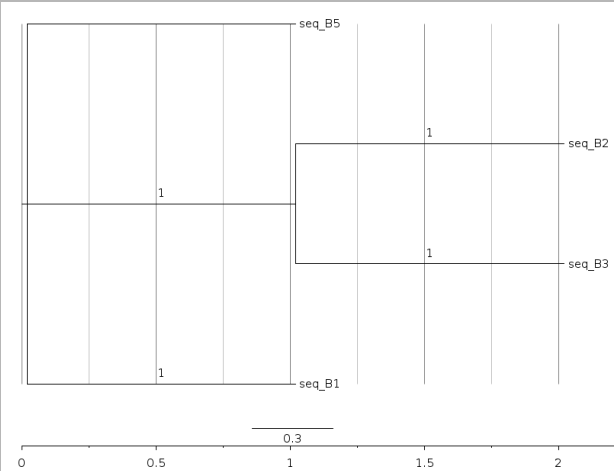
Phylogenetic trees for the given datasets

**Phylogenetic trees for B dataset**

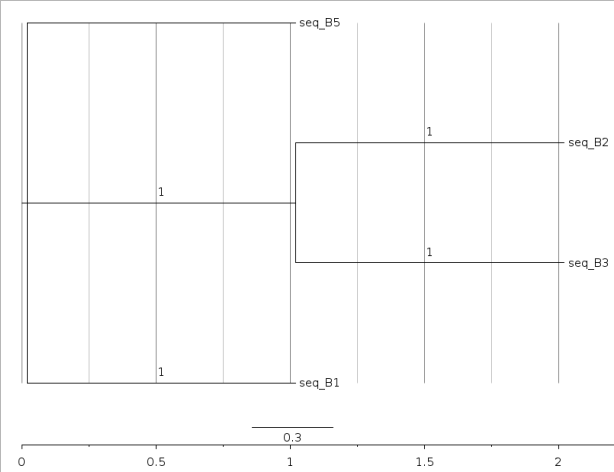
---

<sup>1</sup> Domestic pig. (2018, October 14). Retrieved from [https://en.wikipedia.org/wiki/Domestic\\_pig](https://en.wikipedia.org/wiki/Domestic_pig)

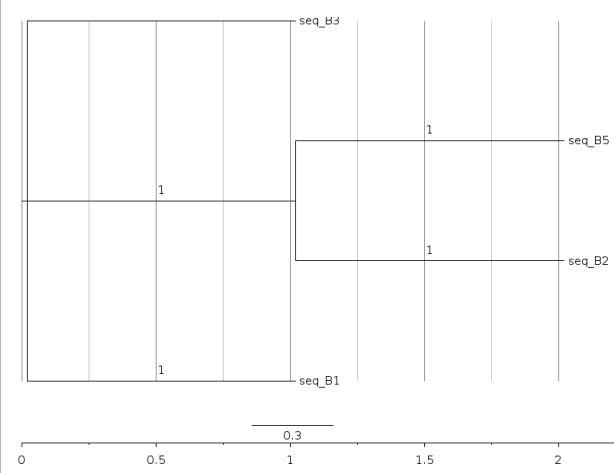
Graph	Figtree Run
	<p data-bbox="1068 478 1172 514">Best tree</p>
	<p data-bbox="1084 951 1156 987">Run 0</p>
	<p data-bbox="1084 1455 1156 1491">Run 1</p>



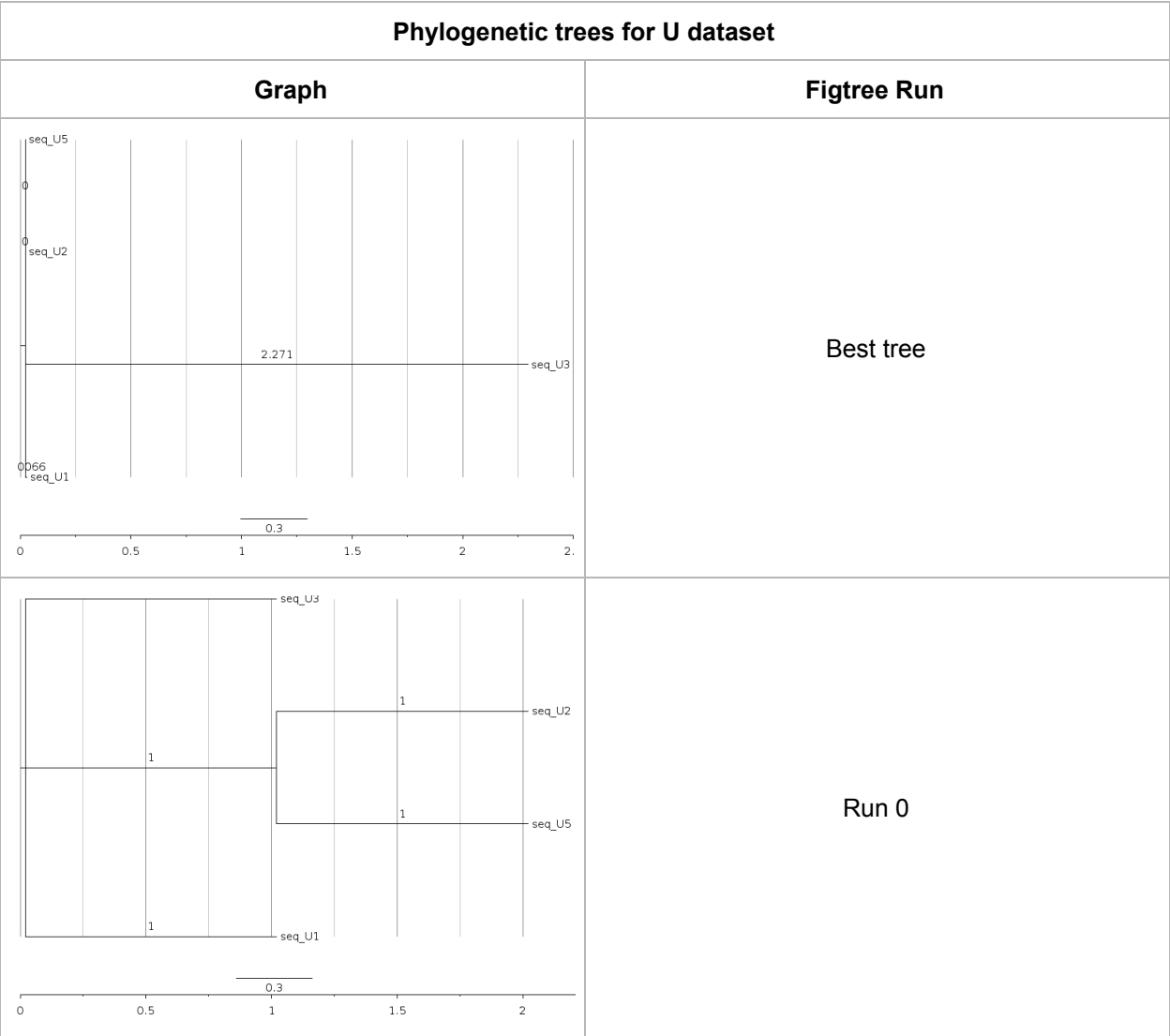
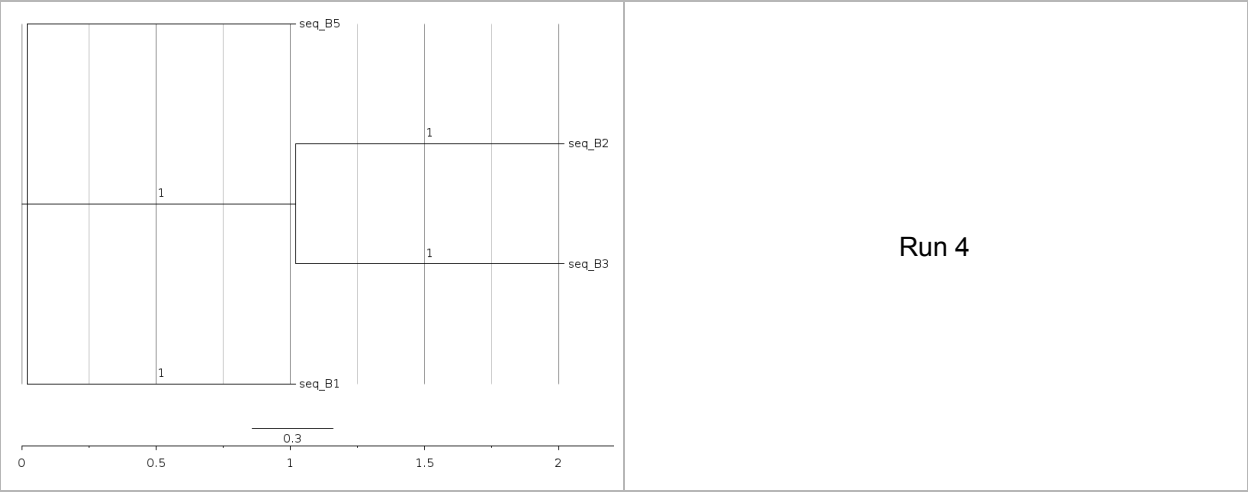
Run 2

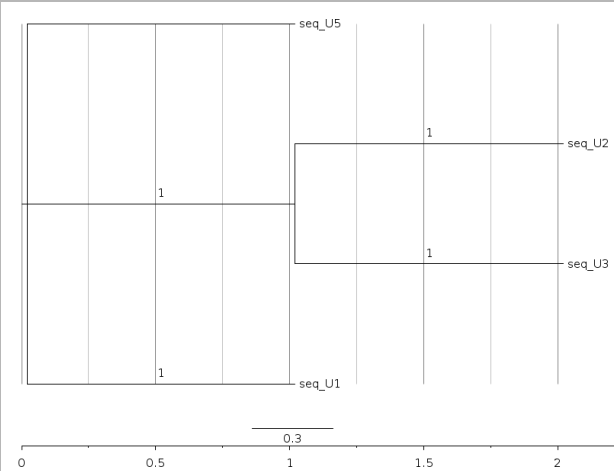


Run 3

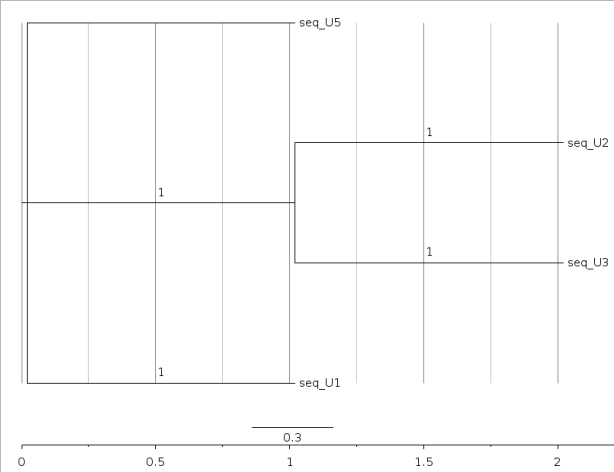


Run 4

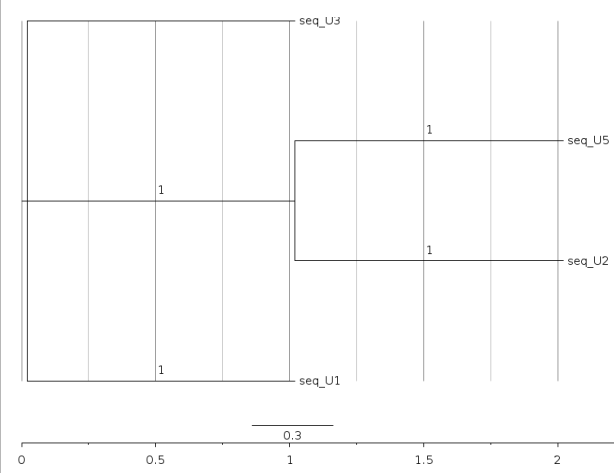




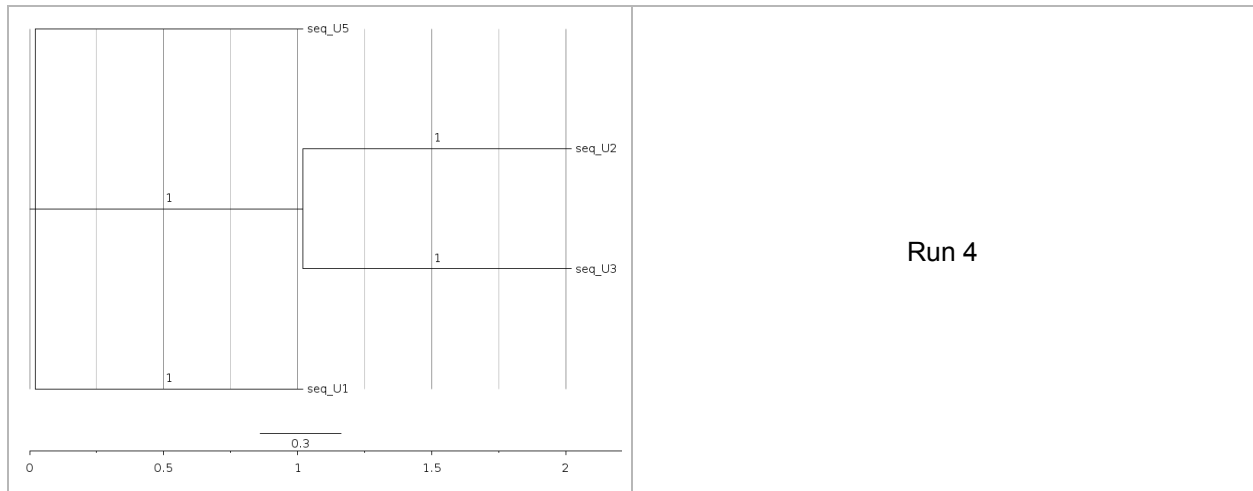
Run 1



Run 2



Run 3



4.

Is there a concurrence between the findings of the BLAST search (i.e. if the meats were properly labelled) and the resulting phylogenetic trees (i.e. were outlier sequences detected in the generated phylogenetic tree)? Expound on your answer.

There is a concurrence between the above BLAST search results and the Figtree-generated phylogenetic trees.

For the B dataset or the alleged “baboy-ramo” dataset, B2, B3, and B5 all appear to be genes of closely related organisms, which the BLAST search identifies as *Sus scrofa* isolate yorkshire. B1 somewhat is somewhat distinct from these (ie. is a separate phylogenetic branch from the above), which concurs with its BLAST search identification as *Sus scrofa* breed Duroc.

Both species were BLAST identified as *Sus scrofa*, which concurs with their visibly closely related phylogenetic branches in the Figtree-generated phylogenetic trees.

This is also the case for all U dataset samples, which were also BLAST identified as *Sus scrofa*, and also appeared to branch from closely related branches in all Figtree-generated phylogenetic trees.

## References

Altschul S.F. et al. (1990) "Basic local alignment search tool".

Domestic pig. (2018, October 14). Retrieved from [https://en.wikipedia.org/wiki/Domestic\\_pig](https://en.wikipedia.org/wiki/Domestic_pig)

Madden T. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. In: McEntyre J, Ostell J, editors. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2002-. Chapter 16. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK21097/>

Rambaut A (2009) FigTree v1.4: Tree Figure Drawing Tool. Available: <http://tree.bio.ed.ac.uk/software/figtree>. Accessed 2018 Oct 21.

Tsukasa Nakamura, Kazunori D Yamada, Kentaro Tomii, Kazutaka Katoh; Parallelization of MAFFT for large-scale multiple sequence alignments, Bioinformatics, Volume 34, Issue 14, 15 July 2018, Pages 2490–2492, <https://doi.org/10.1093/bioinformatics/bty121>