

You must show your work.

1. A portion of a processed pseudogene, believed to be entirely neutral, was sequenced in 10 *Mus musculus* and one *Rattus rattus*. These species diverged 15 MYA and have one generation per year. The sequenced segment was 500 bp long, and 45 differences were observed between species. Within *M. musculus*, 493 sites were invariant and the 7 variable sites had the following distribution.

```

consensus AGCTGGC
Mouse 1   .A.C...
Mouse 2   .A....T
Mouse 3   .A...C.
Mouse 4   G....C.
Mouse 5   .....T
Mouse 6   ....A.T
Mouse 7   G.....
Mouse 8   G....T
Mouse 9   ..T....
Mouse 10  ..T....
    
```

Calculate π and θ and give 95% confidence intervals for each. Estimate the mutation rate per bp for this locus, and estimate the effective population size of mice. Assume that the effective population size of the mouse-rat ancestor was the same as the effective population size of mice today. Calculate Tajima's D and assess whether it is significantly different from 0 (see Table 2 in Tajima 1989 Genetics 123: 585). Is there any evidence for selection in these data?

2. The entire coding region of one gene was sequenced in these same 10 mice and one rat, and 40 non-synonymous and 25 synonymous differences were observed between these species. The coding region consisted of 330 codons, of which 322 were invariant in mice. The 8 variable codons were as follows.

```

Consensus CCC ACT GGG GAT TTA ACT AGT CCC
Mouse 1   ... ..C ...
Mouse 2   ... ..A ...
Mouse 3   ... ..C ...
Mouse 4   ..T ...
Mouse 5   ... ..G ...
Mouse 6   ... ..T ...
Mouse 7   ... ..C ...
Mouse 8   ... ..C ...
Mouse 9   ... ..C ...
Mouse 10  ... ..A ...
    
```

Calculate π and θ and give 95% confidence intervals for each. Calculate Tajima's D and assess whether it is significantly different from 0. Perform a McDonald Kreitman test on these data and assess its statistical significance. Compare patterns of variation at this gene to those seen at the pseudogene above. What do you infer and why? Provide a specific evolutionary scenario that is consistent with the data.