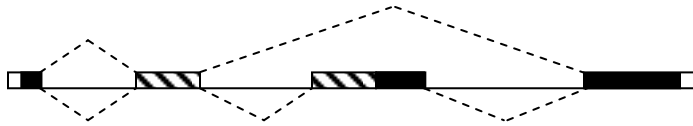




### QUESTION 3

a) Explain what is shown in the schematic of gene X (below) and in your answer include a description of the expected nature of protein X. What evolutionary events could explain the origin and expression of this gene? Propose a second different evolutionary strategy that would achieve the same biological outcome. Assume that the hatched blocks represent duplicated sequences, that open blocks are non-coding, and that 1 cm = 400 bp.



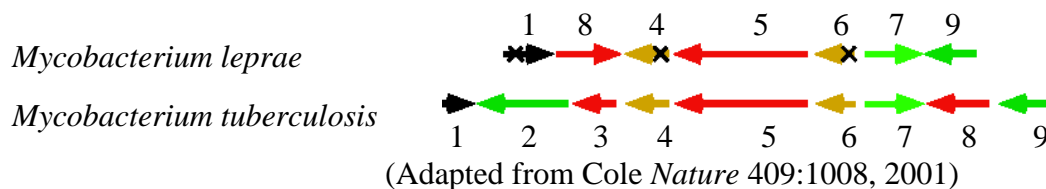
b) Use the data in the above figure to construct two dot matrix plots, one at the level of RNA sequences and the second at the protein level.

### QUESTION 4

Suppose you were able to follow the evolution of a genomic DNA segment (0.1 Mbp in length) over a 20 million year time period. What general types of molecular evolutionary events might be expected? Would you expect any differences in their nature if the DNA were bacterial or from a multicellular eukaryote?

### QUESTION 5

The figure below shows a 20-kb region of the *Mycobacterium tuberculosis* genome (lower) and its counterpart in *Mycobacterium leprae* (upper). Homologous operons (or single genes) are shown by numbered arrows, and the arrows also show gene orientation. Crosses denote pseudogenes.



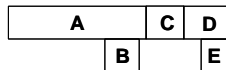
a) Show the dot matrix profile expected when these DNA regions are compared, and briefly give your rationale.

b) Propose a scenario based on the data above which would describe molecular evolutionary events that occurred (in this region of the genome) since the time that these two *Mycobacterium* species shared a common ancestor.

## QUESTION 6

The genomic organization of five genes (A – E) from bacteriophage in the *Microviridae* family is shown below on the left. Sequences for these genes (which in the case of A and D were divided into overlapping and non-overlapping parts) from 30 diverse bacteriophage were used in the analysis shown in the table (adapted from Pavesi *J.Gen.Virol.* 87:1013, 2006). Note that all 5 genes are encoded on the same DNA strand (that is, they are in the same orientation).

- Give an interpretation/explanation of these data.
- Add hypothetical numerical values for C to the table and give your rationale.



A & C: DNA replication proteins  
B & D: capsid morphogenesis proteins  
E: host cell lysis protein

Gene region	$K_s$	$K_n$
A (non-overlapping)	2.07	0.33
A (overlapping gene B)	1.25	0.41
B (overlapping gene A)	0.45	0.56
D (non-overlapping)	1.64	0.29
D (overlapping gene E)	0.81	0.13
E (overlapping gene D)	0.19	0.27
<b>C (non-overlapping)</b>		

### QUESTION 7

This figure (from Campbell et al. *Genome Biol Evol.* 4:289, 2012) has too much data to be used on a test, but it shows interesting features of gene/genome evolution and is useful as a “practice question” for interpretation/explanation. The title of the paper is “Horizontal transfer and death of a fungal secondary metabolic gene cluster.” Tree A shows phylogeny of the *bik6* gene and tree C is based on EF1- $\alpha$ , a housekeeping gene. Note: the *bik1-6* genes are shown in blue and numbers in panel B indicate % nucleotide identity between homologs.

