

## Chapter 13 SELF-TEST

Based on material from chapters 11-13.

1. The Comparative Method for the deduction of secondary structures of RNA molecules relies on the fact that

- A. Substitutions cannot occur in the paired regions of RNA structures.
- B. There are always three hairpin loops in the secondary structure of tRNAs.
- C. Stems in RNA structures do not change their length.
- D. Substitutions in sites that are paired with one another occur in a correlated way.

A – wrong – There are many compensatory substitutions in the paired regions.

B – wrong – Sometimes there are four, but this doesn't answer the question anyway.

C – wrong – Sometimes they change length

D - correct

2. The alignment below is thought to be part of an RNA gene with conserved secondary structure. Which of the four conclusions would you draw? (The species names are abbreviations and are not relevant to the question).

		*		20		*		40	
LAMFLU :	TAAACAAGAA-TAG	TTTGTCTTGT	TAAAGCGTACC	GGAAGGTGCG	CTTA	:			
PETMAR :	TAAACAAGAA-TAG	TTTGTCTTGT	TAAAGCGTACC	GGAAGGTGCG	CTTA	:			
RAJRAD :	TAAGAAAACTTA	GGATAITTTTC	TAAAGTGTACT	GAAAGGTGTA	CTTA	:			
SCYCAN :	TAAGAGAAAGTCT	GAGTATTTCTC	TAAAGTGTACT	GAAAGGTGCA	CTTA	:			
MUSMAN :	TAAGAGAAGATCT	GAGTACTTCCCT	TAAAGTGTACT	GAAAGGTGCA	CTTA	:			
SQUACA :	TAAGGAAAAGTCA	GAGTATTTTCC	TAAAGTGTACT	GAAAGGTGTA	CTTA	:			
CHIMON :	TAAGATAAAAATAG	TATAITTAATC	TAAAGCGTACT	GAAAGGTGCG	CTTA	:			
LEPOCU :	TAAAAAGGACATA	GAGAGTCCTTT	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			
AMICAL :	TAAAAAGGACGCA	GAGAGTCCTTT	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			
ENGJAP :	TAAAGGGGAATG	GAGTGCCCTTC	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			
SARMEL :	TAAAGGGGAAAG	GAGCGTCCCTT	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			
CARAUR :	TAAAGGGGAAAG	GAGTGTCCCTT	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			
POLLOW :	TAAGTAAGAAATAG	GAGCGTCTTAC	TAAAGCGTACC	GGAAGGTGTG	CTTA	:			
POLJAP :	TAAGTAAGAAATAG	GAGCGTCTTAC	TAAAGCGTACC	GGAAGGTGTG	CTTA	:			
GADMOR :	TAAGTAGGAACTAG	GAGTGTCTTGCT	TAAAGCGTACC	GGAAGGTGCG	CTTA	:			
ARCJAP :	TAAGCAGGAAATAG	GAGCGTCTTGCT	TAAAGTGTACC	GGAAGGTGCG	CTTA	:			
PAROLI :	TAAGCAGGGAATAG	GAGTGTCTTGCT	TAAAGTGTACC	GGAAGGTGCA	CTTA	:			

- A. There is evidence for two hairpin loop structures in this region.
- B. This region probably contains the anticodon loop of a tRNA.
- C. This region probably contains a pseudoknot structure.
- D. There is no evidence for secondary structure in this region.

A – correct – The probable secondary structure has been added above and below the alignment. Pairs that are supported by compensatory changes are denoted ( ). Other pairs that are completely conserved and consistent with the supported pairs are denoted [ ].

B – wrong – Anticodon loops almost always have seven unpaired bases in them in which the middle three are the anticodon, and the anticodon sequence is strongly conserved.



- C. The ratio  $d_N/d_S$  can vary between genes because of the varying strength of stabilizing selection acting on the sequences and because of the possibility of positive directional selection acting on some genes.
- D. If the ratio  $d_N/d_S$  is unusually small, this may be an indication that selection on this gene is less stringent than on most genes.

A – wrong – This sentence is complete rubbish!

B – wrong – There is not really any way to link the  $d_N/d_S$  ratio to horizontal transfer.

C – correct

D – wrong – If there were weak stabilizing selection on a gene, this would permit non-synonymous substitutions. Therefore  $d_N/d_S$  would be larger than usual.

6. Which of the following sets of species is listed in order of relatedness to humans from closest to most distantly related?

- A. Baboon, Marmoset, Elephant, Tyrannosaurus, Shark, Amphioxus, Sea urchin.
- B. Chimpanzee, Echidna, Armadillo, Sparrow, Xenopus, Hagfish, Locust.
- C. Haddock, Amphioxus, Locust, Sea urchin, *Saccharomyces cerevisiae*, *Plasmodium falciparum*, *Yersinia pestis*.

D. Armadillo, Cobra, Haddock, Earthworm, *Arabidopsis thaliana*, *Escherichia coli*, *Saccharomyces cerevisiae*

A – correct

B is wrong because the Armadillo is a placental mammal, and is therefore closer to humans than the Echidna, which is a monotreme.

C is wrong because the Sea Urchin is a deuterostome, and is therefore closer to humans than the Locust, which is a protostome.

D is wrong because *S. cerevisiae* is a eukaryote, and is therefore closer to humans than *E. coli*, which is a bacterium.

7. Phylogenetic studies using small sub-unit rRNA showed that...

- A.... there are three principal domains of life known as animals, plants and bacteria.
- B. ...the root of the tree of life is thought to be on the branch separating the bacteria from the other two domains of life.
- C. ...there was an RNA world in the early stages of evolution of life on earth in which RNA carried out the roles played by DNA and proteins in today's organisms.
- D. ... the rRNA genes in chloroplast genomes have a recognizable similarity to those in Cyanobacteria.

A – wrong – The three domains are Archaea, Bacteria, and Eukaryotes.

B – wrong – It is true that the usual view is that the root is on the bacterial branch (although this is still controversial), however, this cannot be shown using SSU rRNA. You need genes that duplicated prior to the divergence of the three domains. This was a tough question, eh?

C – wrong – Many people think that there was an RNA world, but you can't prove it from rRNA phylogenies.

D – correct.

8. Which of these statements about the origin of mitochondria is true?

- A. An organism in which functional copies of genes have been transferred from the mitochondrion to the nucleus is said to have undergone secondary endosymbiosis.
- B. The mitochondrial genomes of all known eukaryotes are thought to share a common origin in a single occurrence of endosymbiosis.
- C. One piece of evidence that mitochondria arose from endosymbiotic bacteria is that the mitochondrial genome of *Trypanosoma brucei* has many organizational features in common with bacterial genomes.
- D. Eukaryotes may be defined as cells possessing mitochondria.

A – wrong – Secondary endosymbiosis is sometimes used to mean that an organism that already contained an endosymbiotic organelle itself became endosymbiotic within a larger cell (which seems to have happened with the evolution of chloroplasts). Secondary endosymbiosis also is used when a second species of endosymbiotic bacterium invades a host that already had one endosymbiotic species (which seems to have happened with some insect endosymbionts). Neither of these things has anything to do with transfer to the nucleus.

B – correct – as far as we know.

C – wrong – *T. brucei* has a very unusual mitochondrial genome that is broken into fragments called minicircles and maxicircles. There are other protists like *Reclinomonas americana* that have mitochondrial genomes that share features with bacteria.

D – wrong – Eukaryotes are cells that possess nuclei, although almost all of them possess mitochondria too.

9. Which of the following statements about bacterial genomes is true?

- A. The smallest known bacterial genome has approximately 500 genes.
- B. The smallest known bacterial genome is approximately 65 kb long.
- C. A single-celled eukaryote like *S. cerevisiae* has roughly 100 times as many genes as a bacterium like *E. coli*.
- D. The genomes of bacteria living as intracellular parasites tend to be larger than those of free-living bacteria because the parasites require complex genetic systems to combat the immune system of their hosts.

A – correct

B – wrong – It is an order of magnitude larger - close to 500 kb

C – wrong – The difference in numbers of genes between the larger bacterial genomes and the smaller eukaryotic genomes is not very great. *E. coli* has roughly 4200 – 5500 genes (depending on the strain) and *S. cerevisiae* has roughly 6300 genes. Several bacteria have more than this.

D – wrong – Intracellular parasites tend to have small genomes. Read Chapter 12!

10. Comparison of bacterial genomes between species indicates that...

- A. ... over 95% of genes on a typical genome have orthologues on every other genome.
- B. ... the more closely related the species, the larger the number of shared orthologous genes they possess.

- C. ... it is very hard to change the order of genes on a genome without disrupting the function.
- D. ... variation in length of bacterial genomes is due principally to variation in length of non-coding regions rather than variation in number of genes.

A – wrong – It is usually far less than this.

B – correct – Phylogenies based on numbers of shared genes appear to be similar to those based on sequence evolution.

C – wrong – Gene order seems to change a lot, suggesting that the order is not crucial for the function in many cases.

D – wrong – It is due principally to variation in the number of genes. However, this statement would be true for some other groups of organisms, such as vertebrates.

11. Two species are found to share a cluster of 8 genes, but the genes are in different orders in the two species. The orders are represented by signed permutations.

species X	1,2,3,4,5,6,7,8
species Y	1,2,-5,-4,-3,8,6,7

The transformation between the two gene orders ....

- A. ...cannot be achieved by inversions alone.
- B. ...can be achieved by one translocation and one inversion.
- C. ...can be achieved by three inversions.
- D. ... requires six separate genome rearrangement events.

A – wrong – You can always do it with inversions alone if you have enough of them.

B – correct – Translocate 8, and invert 3,4,5.

C – wrong – It would need 4 inversions.

D – wrong – Only 2 events, as in B.

12. Which of the following biases in a microarray experiment could be eliminated by a dye flip experiment?

- A. Errors arising from varying numbers of cells of different types in a tissue sample.
- B. Systematic variation of spot size across an array.
- C. Non-specific hybridization of RNA to the probe sequence.
- D. None of these.

D – see Chapter 13

13. Which of the following statements about clustering algorithms is true?

- A. Hierarchical clustering algorithms work by repeatedly dividing a set of points into two until there is only one point remaining in each cluster.
- B. Hierarchical clustering algorithms require the user to pre-specify the desired number of clusters.
- C. K-means clustering minimizes the distances between the points and the centres of the clusters to which they belong.
- D. K-means clustering maximizes the mean distance between points in different clusters.

A – wrong – Hierarchical algorithms work by assembling smaller clusters into larger ones (bottom up), although there are some other methods that cluster from the top down.

B – wrong – The hierarchy works at all levels and can be cut to give any number of clusters. Other algorithms, like K-means, require specification of the number.

C – correct

D - wrong

14. Which of the following statements is true?

A: Cy3 and Cy5 are fluorescent dyes used to visualize the protein spots in polyacrilamide gel electrophoresis.

B: In a 2-dimensional gel, proteins are separated according to size and hydrophobicity.

C: Affymetrix chips require probe sequences several hundred nucleotides long.

D: The probe sequences on microarray chips are usually made of DNA.

A – wrong – These dyes are used to label nucleic acids in microarray experiments.

B – wrong – They are separated according to size and pI.

C – wrong – Usually they are short oligomers

D - correct

15. The yeast two-hybrid system...

A: ... makes use of a transcription factor domain that recruits RNA polymerase to transcribe a reporter gene.

B: ... is a way of locating transcription factor binding sites in the 3' regions of genes.

C: ... makes use of a DNA binding protein from the thermophilic bacterium *Thermus aquaticus*.

D: ... is a way of locating duplicated genes in the yeast genome.

A – correct. See chapter 13.