

BiMM 140

Spring 2003

Quiz 3

Name _____

Key

- (20) What are the basic energy terms used in Zuker's minimum free energy approach to RNA secondary structure prediction? Explain what each means and note which are positive and which are negative.

 - DG (stacking) – interaction between stacked bases in stems - negative
 - DG (hairpin) – hairpins are stems that foldback with only a loop between stem regions – positive
 - DG (bulge) – bulge loops have unstacked bases on only one side – positive
 - DG (internal) – internal loops have unstacked bases on both sides and only 1 stem exiting the loop – positive
 - DG (multibranch) – multibranch loops have multiple branches exiting the loop - positive
- (5) What are “tetraloops” and why are they important?

Tetraloops are hairpin loops with 4-base sequences that occur much more frequently than one would expect at random. Because they are so common, Zuker and others attribute a negative free-energy to these structures.
- (10) Explain the basic principle used to predict RNA structures using a covariance approach?

If two bases are base-paired in a stem, when one mutates, the other must change so as to preserve the base-pairing. This causes the bases at the two positions to covary, or to change in step with each other.
- (20) What are the basic steps used in a progressive multiple alignment approach? Note how the “once a gap, always a gap” principle relates to this procedure.

 - Calculate pairwise distance between sequences (often by DP alignment)
 - Calculate guide tree based on above distance matrix
 - Add sequences to alignment in order indicated by guide tree, when adding sequences the gap penalty is only paid the first time a gap is inserted in an alignment – this is the once a gap, always a gap principle.
- (10) Which approach makes more assumptions about the kind of tree that relates a group of sequences: dynamic programming alignment in multiple dimensions or progressive alignment using neighbor-joining (i.e. CLUSTALW)? Explain.

Both make equal assumptions about the underlying tree. Multi-dimension DP assumes a “star” tree where all sequences diverge equally from a common ancestor, progressive alignment with neighbor joining assumes a hierarchical tree.
- (25) What are five enhancements that CLUSTALW makes over a simple progressive alignment program such as PILEUP?

 - Fast and slow methods to calculate distance matrix
 - Gap penalties adjusted for sequence length and similarity
 - Distance corrected for multiple mutation
 - Neighbor joining tree
 - Scoring matrices depend on distance between sequences
 - Sequences are weighted for similarity during tree construction
 - Position specific gap penalties
 - Composition dependent gap penalties
 - Gapped regions can be omitted in calculating final tree
 - Final tree based on multiple alignment

7. (10) Given a set of trees with different topologies, what criterion does the parsimony method use to determine which one is the best tree?

Parsimony defines the tree that requires the fewest mutations to be the best.