

February 2, 2012

Exam in Applied Computational Molecular Biology Questions

All answers require some (short) argumentation

- 1.** The programs for alignments of large genomic sequences are sometimes specially developed to be optimal for treatment of either closely related species or distant ones. What kind of differences can you expect between parameters of the algorithms for related and distant genomes?
- 2.** Without accurate calculations, try to estimate which of the following two RNA fragments could have more stable secondary structure (lower free energy)?

- (a) ACCCAAAGGGAAACCCAAAAGGGAAACCCAAAAGGGAAACCCAAAAGGGAA
(b) ACCCAAAGGGGAACCCAAAAGGGGAACCCAAAAGGGGA
(c) ACCCCCAAAGGGGGGAACCCCAAAGGGGGGA

- 3.** A bioinformatician X has designed a program for RNA pseudoknot prediction. In the RNA fragment, given below, the program predicted the pseudoknot shown in the bracket view. However, a molecular biologist Y, working on this RNA, says that this is wrong prediction because a different functional secondary structure is found experimentally. Try to suggest an alternative structure for this fragment (give a bracket view):

```
CCCUCGUUCCGAGGGUCAUCGGA   RNA
(((.....[[[[[[]))...]]]])   pseudoknot prediction
??? different structure
```

- 4.** Below a fragment of multiple alignment of four sequences is given. On the basis of variation in this alignment, try to suggest a conserved consensus structure in this region.

```
AGGUCUGGAGCAGCUGGAGCAGCUGCGAAAGGA seq1
AGGUCUGGGGCAGCUGGAGCAGCUGCGAAAGGA seq2
AGGUCUGGAGCUGCAGGCGCUGCAGCCAAAGGA seq3
AGGUCUGGAGCUGCAGGCGCUGCAGCCAAAGGA seq4
***** ** ** ** ** ** ** ** ** ** ** ** ** ** **   conserved nucleotides
??? consensus structure
```

- 5.** Programs for searching the targets of miRNA binding tend to yield many false-positive results. What are the reasons for that? What are the ways to improve the performance of these algorithms?

- 6.** Assume you need to design an RNA fragment that would fold into the following structure (a hairpin and a stem-loop structure):

```
(((((.....)))).....((((.....((((.....)))).....))))
```

Which of the following sequences is the most suitable:

- (a) CCCCAAAGGGGAAAACCCCAAAGGGGAAAACCCCAAAGGGG
(b) CCGCAAAGCGGAAAACCGCAAAGCGGAAAACCGCAAAGCGG
(c) CCCCAAAGGGGAAAACCGCAAAGGGGAAAACCGCAAAGCGG
(d) CGCAAAGGCGAAAACGCCCAAAGGCGAAAACGCCCAAAGGCG

Try to suggest another construct that would fold into the same structure.

- 7.** The RNA sequence given below can fold into two rather stable alternative secondary structures. What are they? Try to estimate which of them would fold faster (with higher folding rate constant).

```
AGGGCGAAAGCCCGAAAGGGCGAAAGCCCA
```

