

Namn och Personnummer

Instruktioner

Tillåtna hjälpmedel: papper och penna.

Varje fråga skall besvaras inom det befintliga utrymmet, om ni behöver mer utrymme skriv på baksidan.

Glöm inte att skriva namn och personnummer på alla papper. Om dessa saknas kan det hända att dina svar inte tas med i din totalpoäng!!

Du kan svara på alla frågor på svenska eller engelska.

För Betyg G krävs 15 poäng (50%), Betyg VG 22.5 poäng (75%).

Instructions

Allowed tools: Paper and Pen and Dictionary.

Each question should be answered within the provided space, if you need more space write on the back of the paper.

Namn och Personnummer

1. Beskriv hur s.k. phylogenetic profiles kan användas för att förutsäga protein-proteininteraktioner (3p).

Namn och Personnummer

2. Var hittar man följande aminosyror i strukturerna av de kända membranproteinerna av 'helix bundle'-typ: (1) Trp, Tyr; (2) Lys, Arg (2p)?

Namn och Personnummer

3. Beskriv kortfattat hur PSI-BLAST fungerar (3p) ?

Namn och Personnummer

4. Du har tre proteiner, A B och C som är 50, 100 och 200 residuer långa. Efter upplinjering av dessa proteiner ser du att A,B är 30% identiska, B, C är 45 % identiska och A,C 25% identiska. Vilka av dessa proteiner är säkerligen homologa (2p) ? Hur kan du göra ett bättre test för att avgöra detta (1p) ?

Namn och Personnummer

5. Du har hittat en ny avlägsen homolog (A) till en TIM barrel och en annan homolog (b) till Ribosomal protein S19. Vad tror du att du kan säga om funktionen av A och B (2p) ?

Namn och Personnummer

6. Om man använder en domändatabas såsom Pfam (eller SCOP) så ser man att vissa domäner är mycket vanligare än andra. Vad kallas denna fördelning (1p)? Beskriv en teori hur denna fördelning kan ha uppkommit (2p) ?

Namn och Personnummer

7. *Background:* The gene family tenT-A comprise three homologs in human (h1, h2 and h3), four homologs in gorilla (g1, g2, g3 and g4) and two homologs in orang-utan (o1 and o2). The reconstructed gene tree for this, its reconciliation to the corresponding species tree and the identification of speciation, duplication, loss and lateral transfer events in the gene tree is shown below:

Questions: Name which of the three types of homology apply to the following pairs of genes (3p):

- (a) g3 - h1
- (b) g1 - g2
- (c) h3 - o2
- (d) o1 - g3
- (e) o2 - g4
- (f) g3 - h2

Namn och Personnummer

8. Give two different reasons why it is harder to predict the presence of a gene in a genomic DNA sequence in eukaryotes compared to prokaryotes (2p).

Namn och Personnummer

9. Why is normalization an important issue in the analysis of microarray data (3p)?

Namn och Personnummer

10. Fold recognition (FR) techniques attempt to predict protein structure by looking at the "goodness of fit" of the query sequence (Q) against a number of database proteins (D1, D2 ... Dn) whose 3D structures are known. What type of structural information can be used to calculate this "goodness of fit" (3p) ?

Namn och Personnummer

11. In this question, secondary structure prediction refers to the prediction of each amino acid residue in a globular protein into three states: helix, strand and coil.
 - (a) Explain a simple way to predict secondary structure from a single amino acid sequence using simple statistics (such as that used by Chou and Fasman in the early 1970s). (2p)
 - (b) Why do multiple sequence alignments of family members give higher quality input information for secondary structure prediction? (1p)