

INF4350 Høst 2009 Deleksamen 2

Generell informasjon

- Denne deleksamen skal løses individuelt.
- Den teller 15% på den endelige karakteren.
- Vedlagte erklæring ”Krav til innleverte oppgaver ved Institutt for informatikk (Ifi)” skal leses og aksepteres.
- Skriv en kort rapport med beskrivelse av din fremgangsmåte for å løse oppgavene, et sammendrag av resultatene underveis og svar på spørsmålene.
- Kontakt faglærer om noe er uklart eller hvis du står helt fast.
- Rapporten må leveres senest fredag 27. november 2009 kl 1700 per epost til torognes@ifi.uio.no i form av ett enkelt PDF-dokument.

Oppgaver

- 1) Finn proteinsekvensen i databasen RefSeq til det humane genet med symbol NTHL1. Proteinets skal være 312 aminosyrer langt. Vis sekvensen i FASTA-format. (Formatter sekvensene med Courier New eller en annen skrifttype med fast bredde for å få det til å se bra ut.)
- 2) (a) Hva er de latinske navnene på modellorganismene bananflue, bie, kylling, mais, mus, sebrafisk og storfe? (b) Identifiser sannsynlige ortologer til det humane NTHL1-proteinet hos disse modellorganismene, samt *Mycobacterium tuberculosis* H37Rv. Forklar hvordan du fant disse ortologene. (c) Hent ut alle 8 proteinsekvensene fra RefSeq-databasen. Vis sekvensene i FASTA-format.
- 3) Samle alle de totalt 9 FASTA-formatterte proteinsekvensene i samme tekstfil. Rediger de første linjene med sekvensbeskrivelsen (linjene som begynner med ">") slik at du bare beholder navnet på organismen. Erstatt mellomrom i navnene med understrek (F.eks. >Homo_sapiens). Start programmet Jalview version 2.4.0b2 (hentes fra www.jalview.org). Hent inn sekvensene ved å velge Input Alignment from File eller Input Alignment from Textbox på File-menyen. Lag en flersekvenssammenstilling av sekvensene med Muscle-algoritmen som du kan kjøre fra WebService-menyen i Jalview. Slå av annoteringer (View – Show Annotations). Slå på linjeombrekking (Format – Wrap). Velg ClustalX-farger (Colour– ClustalX). Gjør vinduet så bredt at du ser ca 60 aminosyrer i bredden. Eksporter sammenstillingen som et bilde (png) og inkluder det i rapporten.
- 4) Lag også en sammenstilling ved hjelp av ClustalW-algoritmen på samme måte som over. Inkluder også denne i rapporten.
- 5) Lag fylogenetiske trær fra sekvensene i Jalview basert på ClustalW-sammenstillingen. Benytt alle de fire tilgjengelige metodene under Calculate Tree på Calculate-menyen, dvs Average Distance (UPGMA) og Neighbour Joining (NJ), begge både med % Identity og BLOSUM62 som utgangspunkt for distansematrixene. (a) Vis de fire trærne. Angi hvilken metode som er benyttet for hvert tre. (b) Hvilket av trærne rimer best med den

allmenne oppfatningen av hvordan evolusjonen har foregått? Hva er det som ikke stemmer i de andre trærne?

- 6) I den C-terminale enden av proteinene er det et karakteristisk motiv som går igjen i alle sekvensene. Det består av 4 cysteiner og noen andre aminosyrer mellom dem. (a) Hva kalles dette motivet? (b) Har ClustalW og Muscle klart å sammenstille dette motivet riktig i alle sekvensene, dvs klart å få de korresponderende cystein-aminosyrene i samme kolonne?
- 7) Hent ut sekvensene (i FASTA-format) fra dette motivet som inkluderer de 4 cysteinene fra de 8 sekvensene. (Dette kan gjøres ved enkelt å markere samtlige aminosyrer og deretter velge Copy fra Edit-menyen.) Vis sekvensene.
- 8) Bruk programmet Weblogo 3 (<http://weblogo.threeplusone.com/>) til å lage en figur som illustrerer dette motivet. Vis figuren.
- 9) Det finnes en struktur av det ortologe proteinet fra bakterien *Escherichia coli* i PDB-databasen. (a) Hvordan kan man finne dette? (b) Hvilken PDB ID har denne strukturen? (c) Inkluder et bilde av denne strukturen i rapporten.



Institutt for informatikk

Krav til innleverte oppgaver ved Institutt for informatikk

Ved alle pålagte innleveringer av oppgaver ved Ifi – enten det dreier seg om obligatoriske oppgaver, hjemmeeksamen eller annet – forventes det at arbeidet er et resultat av studentens egen innsats. Å utgi andres arbeid for sitt eget er uetisk og kan medføre sterke reaksjoner fra Ifis side.

Derfor gjelder følgende:

1. Hvis du tar med tekst, programkode, illustrasjoner og annet som andre har laget, må du tydelig merke det og angi hvor det kommer fra.
2. Det er greit å få hint om hvorledes en oppgave kan løses, men dette skal eventuelt brukes som grunnlag for egen løsning og ikke kopieres uendret inn.
3. Kursledelsen kan innkalle studenter til samtale om deres innlevering.

Gruppearbeid

I noen kurs skal det leveres gruppearbeid. Ifi krever da at alle medlemmer av gruppen kan gjøre rede for hovedtrekkene i det innleverte arbeidet. Dessuten må alle ha utført en rimelig del av det hele, og kunne identifisere og svare i detalj for sin del.

Samarbeid

Reglene om kopiering betyr ikke at Ifi fraråder samarbeid – tvert imot, Ifi oppfordrer studentene til å utveksle faglige erfaringer om det meste. Men det kreves som nevnt at man kan stå inne for det som leveres.

Hvis du er i tvil om hva som er lovlig samarbeid, kan du kontakte gruppelærer eller faglærer.

www.ifi.uio.no/studinf/skjemaer/erklaring.pdf

27. jan. 2004