



Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Oswaldo Cruz
PG-BCS

Prova de seleção para o doutorado em Biologia Computacional e Sistemas

Artigo: Raab & Tvaroska, 2010 - The binding properties of the H5N1 influenza virus neuraminidase as inferred from molecular modeling.

Questões

1) Explique porque no contexto do trabalho desenvolvido por Raab & Tvaroska o uso da modelagem molecular como alternativa aos métodos experimentais foi justificado.

2) Métodos de docking proteína-ligante são amplamente utilizados em modelagem molecular. Sobre docking molecular e tendo como base o artigo escolhido, responda à seguinte questão:

- Descreva detalhadamente como os resultados obtidos a partir de um procedimento de docking molecular podem ser avaliados, considerando a indisponibilidade de uma estrutura experimental do complexo em estudo. Justifique o critério adotado (ou, os critérios adotados) em cada etapa de sua resposta.

3) Em função dos dados da figura 7, poderia se afirmar que o modo de ligação do ácido siálico, determinado pelo experimento de docking, é o correto? Justifique a resposta.

4) No presente trabalho, as simulações por DM foram realizadas utilizando o campo de forças ff99SB para a descrição da proteína, Glycam06 para carboidratos junto com o TIP3P para o modelo de solvente. Se a simulação tivesse sido realizada utilizando o campo de força Gromos96 para a descrever a proteína, OPLSAA para a caracterização de carboidratos, e modelo SPC/E para o solvente, dizer se se teriam obtido resultados equivalentes e confiáveis. Justificar.



Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Oswaldo Cruz
PG-BCS

Prova de seleção para o doutorado em Biologia Computacional e Sistemas

Artigo: Maher et al., 2010 - Transcriptome sequencing to detect gene fusions in cancer

Questões

- 1) Discuta quais são as vantagens e desvantagens das diferentes plataformas de seqüenciadores de DNA de segunda geração (NGS) para a análise do transcriptoma. Aponte ao menos uma vantagem de usar mais de uma tecnologia de NGS para a análise do transcriptoma de uma amostra tumoral ou tipo celular.
- 2) A definição de gene é um tema em discussão na literatura. Neste sentido, proponha uma definição para gene e sugira uma explicação para Maher e colaboradores (2009) terem detectado sequencias oriundas do transcriptoma contendo éxons de dois genes presentes no mesmo cromossomo sem a presença de rearranjos genômicos.
- 3) Defina e discuta as vantagens e desvantagens do uso de técnicas de normalização e subtração de cDNA para estudar o transcriptoma.
- 4) Discuta o algoritmo que Maher e colaboradores (2009) utilizaram para definir um transcrito quimérico usando dados de seqüenciadores de DNA de segunda geração. Você teria alguma sugestão para aprimorar a abordagem dos autores?



Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Instituto Oswaldo Cruz
PG-BCS

Prova de seleção para o doutorado em Biologia Computacional e Sistemas

Artigo: Coulet et al., 2010 - Using text to built semantic network for pharmacogenomics.

Questões:

1) Coulet e colaboradores apresentam um novo método que usa a estrutura sintática das sentenças para extrair relações biomédicas de um texto. Qual a utilidade em organizar entidades em ontologias cujas relações semânticas sejam comuns e resumidas numa rede semântica normalizada?

2) Quais as dificuldades encontradas na extração das regras para mineração de texto em farmacogenômica?

3) Em linhas gerais, quais são os quatro passos do método proposto pelos autores e o que cada um deles se propõe?

4) O método proposto por Coulet e colaboradores permite criar sumários de relações semanticamente ricas entre entidades-chave (palavras chaves) reconhecidas e classificadas como genes, drogas e fenótipos, e também entre entidades modificadoras destas entidades-chave, sem criar uma enumeração a priori de tipos de relações.

- i. Quais as classes de palavras-chave, sementes ou entidades farmacogenômicas usadas para extrair as relações a serem grupadas na ontologia?
- ii. Qual o par de classes de entidades (sementes) que não foi considerada na seleção das frases (sentenças) de interesse, e por quê?