

TAREFA II de BIOINFORMATICA

Entregar para biodados@gmail.com usando como assunto = “TAREFA2” até 9/7 meio dia

- De acordo com o número constante na lista de chamada, trabalhe com a via determinada abaixo, caso não haja hit em humanos, escolha uma do final (>45)

1. Glycolysis / Gluconeogenesis
2. Citrate cycle (TCA cycle)
3. Pentose phosphate pathway
4. Pentose and glucuronate interconversions
5. Fructose and mannose metabolism
6. Galactose metabolism
7. Ascorbate and aldarate metabolism
8. Starch and sucrose metabolism
9. Amino sugar and nucleotide sugar metabolism
10. Pyruvate metabolism
11. Glyoxylate and dicarboxylate metabolism
12. Propanoate metabolism
13. Butanoate metabolism
14. C5-Branched dibasic acid metabolism
15. Inositol phosphate metabolism
16. Oxidative phosphorylation
17. Photosynthesis
18. Photosynthesis - antenna proteins
19. Carbon fixation in photosynthetic organisms
20. Reductive carboxylate cycle in photosynthetic bacteria
21. Methane metabolism Major update!
22. Nitrogen metabolism
23. Sulfur metabolism
24. Fatty acid biosynthesis
25. Fatty acid elongation in mitochondria
26. Fatty acid metabolism
27. Synthesis and degradation of ketone bodies
28. Steroid biosynthesis
29. Primary bile acid biosynthesis
30. Secondary bile acid biosynthesis
31. Steroid hormone biosynthesis
32. Glycerolipid metabolism
33. Glycerophospholipid metabolism
34. Ether lipid metabolism
35. Sphingolipid metabolism
36. Arachidonic acid metabolism
37. Linoleic acid metabolism
38. alpha-Linolenic acid metabolism
39. Biosynthesis of unsaturated fatty acidsAlanine, aspartate and glutamate metabolism
40. Glycine, serine and threonine metabolism
41. Cysteine and methionine metabolism
42. Valine, leucine and isoleucine degradation
43. Valine, leucine and isoleucine biosynthesis
44. Lysine biosynthesis
45. Lysine degradation
46. Arginine and proline metabolism
47. Histidine metabolism
48. Tyrosine metabolism
49. Phenylalanine metabolism
50. Tryptophan metabolism
51. Phenylalanine, tyrosine and tryptophan biosynthesis

- Acesse o site do Kegg (<http://www.genome.jp/kegg/pathway.html>)
- Clique na via determinada acima

- Mude, no menu persiana, “Reference pathway” para “Homo sapiens (human)” e aperte o botão [Go]
- Escolha 5 enzimas marcadas em verde (presentes em humanos)
- Para cada enzima escolhida, clique sobre o retângulo verde
- Procure a sequência nucleotídica. Próximo a ela há um botão [NT seq], clique para obter a sequência e copie em um documento em seu computador local.
- Copie em outro documento a descrição da sequência e deixe somente o identificador Kegg na sequência. Exemplo:

>hsa:7086 TKT, FLJ34765, TK, TKT1; transketolase (EC:2.2.1.1); K00615 transketolase [EC:2.2.1.1] (N)

Deixe na sequência apenas

>has: 7086

Em outro documento guarde as informações associadas:

hsa:7086 = TKT, FLJ34765, TK, TKT1; transketolase (EC:2.2.1.1); K00615 transketolase [EC:2.2.1.1] (N)

- Note que algumas enzimas têm várias isoformas, copie todas!
- Quando terminar de coletar as sequências, entre em sua conta na maracatu
- Crie um diretório prova [mkdir tarefa2]
- Entre no diretório [cd tarefa2]
- Crie um documento seqs com o programa vi [vi seqs]
- Entre no modo de inserir [i]
- Copie as sequências [control c] e cole no terminal [botão direito]
- Salve [Esc, :, x!]
- Confira [more seqs] [q]
- Por meio de alinhamento com blastn vamos contar quantos hits cada sequência tem na amostra de reads de 454:
- megablast -i seqs -d /home/bacharelado/blast_aula/454data/tumor.seq -D 3 -F F -a 4 -p 96 -s 100 -o saída
- Verifique a saída [more saída] [q]
- Tire as linhas com jogo da velha e salve como boa:
- [cat saída | grep -v “#” > boa]
- Contemos quantas hits cada query obteve:
- [cat boa | awk ‘{print \$1}’ | sort | uniq -c | more]

RELATÓRIO

1. Reporte o resultado do BLAST
2. Analise o resultado e comente.
3. Como seria possível analisar todas as vias do Kegg de uma única vez?