

Práticas de Bioinformática – Análise de Sequência de DNA de dinossauros

Introdução

Leia a primeira página do trabalho de Boguski (1992), fornecida abaixo. O trabalho é sobre o livro “Parque dos Dinossauro” de Micheal Crichton. A sequência de DNA que aparece no trabalho é reproduzida abaixo como “Jurassic Park Dino DNA from the book Jurassic Park”.

Posteriormente um segundo livro foi publicado por Michael Crichton, chamado “Mundo Perdido” e uma nova sequência de DNA foi mostrada e é reproduzida abaixo como “Lost World Dino DNA from the book The Lost World”.

Conjunto de dados

>Jurassic Park Dino DNA from the book Jurassic Park

```
gcgttgctggcggttttccatagctccgccccctgacgagcatcacaatacgacgcgggtggcgaaacccgacaggactataaagataccaggcggtt
ccccctggaagctccctcggttccgacctgcccgttacgggatactgtccgcctttctcccttcgggaagcggtgctcacgctgtaggtatctcagtt
cgggtgtaggtcgttcccaagctgggctgtgtccggtcagcccgacctgctgccttatccggtaactatcgtctgagtcacacccggtaaagtagga
cagggtccggcagcgctctgggtcattttcggcgaggaccgtttcgtggagatcgccctgtcgttgcggtattcggaaatcttcacgcccctcgtcaag
cttcgtcactccaaacgtttcggcgagaagcaggccattatcgccggcatggcgccgacgcgctgggctggcggttcgcgacgcgaggtggtatggc
cttccccattatgattcttctcgttccggcgcccgctgtgcaggccatgctgtccaggcaggtagatgacgaccatcaggacagcttcaacggctcttac
cagcctaacttcgatcactggaccgctgatcgtcacggcgatttatgccgaagtgcagggtggcgaaacccgacaaggactataaagataccaggcggtt
ccccgggaagcgctctcctgttccgacctgcccgttacgggatactgtccgcctttctcccttcgggctttctcattgctcacgctgtaggtatctcagttcgg
tgtaggtcgttcccaagctgacgaacccccgttcagcccgacctgctgccttatccggtaactatcgtctgagtcacacacgacttaacgggttgg
catggattgtaggcgccgacctatactgtgtgctcccccggtgcatggagccgggcccactcgacctgaatggaagccggcgccgacctcgtcaac
ggccaagaattggagccaatcaattcttgcggagaactgtgaatgcgcaaacaccccttggccatcgctccgcatctccagcagccgcacgcggcg
catctcgggcagcggttgggtcctgcgcatgctgtgtagcctgtcgttgaggaccggctaggctggcggggttgcctactatgaatcaccgatacgcg
agcgaacgtgaagcgactcgtcgtgcaaacgtctgcgacctatgaatggtcttcggtttcgttaagtctggaacgcggaagtcagcgccctg
```

>Lost World Dino DNA from the book The Lost World

```
gaattccggaagcagagcaagataagtcctggcatcagatacagttggagataaggacggacgtgtggcagctcccgacaggattcactggaagtgc
attacatcccatgggagccatggagttcgtggcgctggggggcgcgatgcgggctccccactccgttccctgatgaagccggagccttctggggc
tggggggggcgagaggacggaggcgggggggctgtggtcctctacccccctcaggccgctgttccctggtgcctgggagacacgggtacttt
gggggacccccagtggtgcccgcgcccaaatggagccccccactacgtgagctgctgcaaccccccggggagccccccccatccctcc
tccgggcccctactgccactcagcagcgggccccccacctgcgagggcctgagtgctgcatggcaggaagaactgcggagcgacggcaacgcgg
ctgtggcgccgggacggcaccgggcttacctgtgcaactgggctcagcctgcgggctctaccaccgctcaacggccagaacccgcccgtcatcg
ccccaaaagcgcttgcgggtgagtaagcgcgagggcagtgctgcagccacgagcgtgaaaactgcagacatccaccaccactctgtggcgtgc
agccccatgggggaccccgctgcaacaacattcacgcctgcggcctctactacaaactgcaccaagtgaaccgccccctcacgatgcgcaagacggga
atccaaacccgaaacgcgaaagtcttccaaagggtaaaaagcgcgccccccgggggggaaacccctccgcccacgggagggggcgctcct
atgggggggaggggggagccctctatgcccccccgccgccccccggcgccgccccctcaaagcgagcgtctgtacgctctcggccccgtgg
tcttttgggcccattttcgtccctttgaaactccggaggggttttggggggggggcggggggttacaggccccccggggctgagccccgagattaa
ataataactctgacgtgggcaagtgggcttctgagaagacagtgaacataataatttcacctcggaattgcagagggctgactccactttggacaca
acagggtactcgttaggaccagataagcatttctcctgactgaaaaagaaaggattatctgttcttctgctgacaaatccctgtgaaaggtaaaa
gtcggacacagcaatcgattatttctgcctgtgtgaattactgtgaatattgaaatatatatatatatatctgtatagaacagcctcggaggcggc
atggaccagcgtagatcatgctgatttgactgcccgaattc
```

Referências

Bioinformatics Tutorial – “Jurassic Park” Dino-DNA analysis. <http://www.inf.fu-berlin.de/lehre/WS05/aldabi/aufgabe5_12.html>, acesso em maio de 2015.

Boguski, M. S. A molecular biologist visits Jurassic Park. BioTechniques, v. 12, n. 5, p. 668-669, 1992.

Protocolo

- 1 Proponha uma forma de analisar as sequências de DNA para confirmar se realmente são de origem de um dinossauro extinto?
- 2 Estas são mesmo sequências de dinossauro?

- 2.1 Se são sequências de dinossauro, qual é a espécie?
- 2.2 Se não são sequências de dinossauro, de onde as sequências se originaram?
 - 2.2.1 Alguma destas sequências se aproxima mais da sequência de um dinossauro?
- 2.3 Sendo ou não sequências de dinossauro, estas são sequências de algum gene?
 - 2.3.1 Se sim, qual é este gene?
 - 2.3.2 Este gene codifica para uma proteína? Qual?
- 2.4 Para a sequência “Jurassic Park Dino DNA from the book Jurassic Park”, compare seus resultados com aqueles obtidos por Boguski (1992).
- 2.5 5. A sequência “Lost World Dino DNA from the book The Lost World” possui uma mensagem encriptada. Você é capaz de descobrir qual?