

Task3: Biostrings

Alex-Alex-Helena

17 oktober, 2021

Secuencia DNA

El fichero fasta que se lee es *./Fastas/NC_000001.11.fasta*.

Algunas características

Nombre de la secuencia:

NC_000001.11:c229434094-229431245 Homo sapiens chromosome 1, GRCh38.p13 Primary Assembly

Longitud de la secuencia: 2850 pares de bases.

Extracción de un segmento de la secuencia

Se extraen desde las posiciones iniciales 980, 1215, 1664, 1910, 2193, 2453 hasta 1108, 1539, 1825, 2101, 2374, 2596, respectivamente.

La secuencia extraída es:

```
ATGTGCGACGAAGACGAGACCACCGCCCTCGTGTGCGACAATGGCTCCGGCCTGGTGAAAGCCGGCTTCGCCGGGGATGA
CGCCCCTAGGGCCGTGTTCCCGTCCATCGTGGGCCGCCCGGACACCAG, GCGGTCATGGTCGGTATGGGTCAGAAAGA
TTCCTACGTGGGCGACGAGGCTCAGAGCAAGAGAGGTATCCTGACCCTGAAGTACCCTATCGAGCACGGCATCATACCA
ACTGGGATGACATGGAGAAGATCTGGCACCACACCTTCTACAACGAGCTTCGCGTGGCTCCCGAGGAGCACCCACCCTG
CTCACCGAGGCCCCCTCAATCCCAAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACGTGCCCGC
CATGTACGTGGCCATCCAGGCCGTGCTGTCCCTCTACGCCTCCGGCAGGACCACCG, GCATCGTGTGGACTCCGGCGA
CGGCGTCACCCACAACGTGCCATTTATGAGGGCTACGCGTGCCGCACGCCATCATGCGCCTGGACCTGGCGGGCCGCG
ATCTACCGACTACCTGATGAAGATCCTCACTGAGCGTGGCTACTCCTTCGTGACCACAG, CTGAGCGGAGATCGTGC
GCGACATCAAGGAGAAGCTGTGCTACGTGGCCCTGGACTTCGAGAACGAGATGGCGACGGCCGCTCCTCCTCCTCCCTG
GAAAAGAGCTACGAGCTGCCAGACGGGCAGGTCATCACCATCGGCAACGAGCGCTTCGCTGCCCGAGACGCTCTTCCA
GCCCTCCTTCATCG, GTATGGAGTCGGCGGGCATTACGAGACCACCTACAACAGCATCATGAAGTGTGACATCGACAT
CAGGAAGGACCTGTATGCCAACAACGTCATGTCGGGGGGCACCAGATGTACCCTGGGATCGCTGACCGCATGCAGAAAG
AGATCACCGCGCTGGCACCAGCACCATGAAGATCAAG, ATCATCGCCCCCGGAGCGCAAATACTCGGTGTGGATCG
GCGGCTCCATCCTGGCCTCGCTGTCCACCTTCCAGCAGATGTGGATCACCAAGCAGGAGTACGACGAGGCCGGCCCTTCC
ATCGTCCACCGCAAATGCTTCTAG
```

Este segmento tiene longitud: 1144.

Traducción a Proteína

La traducción de la secuencia produce la siguiente proteína:

```
MCDEDETTALVCDNGSLVKAGFAGDDAPRAVFPISIVGRPRHQGVMMVGMGQKDSYVGDEAQS KRGILTLKYP I EHG IITN
WDDMEKIWHHTFYNELRVAP EHP TLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGR TTGIVLDSGDG
```

VTHNVPIYEGYALPHAIMRLDLAGRDLTDYLMKILTERGYSFVTTAEREIVRDIKEKLCYVALDFENEMATAASSSSLEK
SYELPDGQVITIGNERFRCPETLFQPSFIGMESAGIHETTYNSIMKCDIDIRKDLYANNVMSGGTTMYPGIADRMQKEIT
ALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWITKQEYDEAGPSIVHRKCF

Esta proteína tiene 377 aminoácidos.

Lectura de una proteína

El fichero fasta que se lee es *./Fastas/NP_001091.1.fasta*.

Algunas características

Nombre de la proteína:

Actin, alpha skeletal muscle

Longitud de la secuencia: 377 aminoácidos.

Comprobación que la secuencia traducida corresponde a la secuencia leída

AVISO: La secuencia traducida corresponde a la secuencia leída.