

TRABAJO BASH

#EXTRAE EL NÚMERO DE SECUENCIAS QUE HAY EN EL ARCHIVO

#EXTRAE EL IDENTIFICADOR DE LAS SECUENCIAS

#EXTRAE LAS SUBSECUENCIAS DE TU ARCHIVO MULTIFASTA QUE EMPIECEN POR ATGC y QUE ACABEN POR GTAA. ¿HAY ALGUNA?

```
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -c ">" HBBmultifasta.txt
7
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -E -o "^[A-Z]{2}\.[0-9]" HBBmultifasta.txt
>NM_000518.5
>NM_001164429.1
>NM_002822127.5
>NM_032240524.1
>NM_037840313.1
>NM_054440830.1
>NM_059322707.1
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -E -o "ATGC.*GTAA" HBBmultifasta.txt
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep "ATGC.*GTAA" HBBmultifasta.txt
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -c "ATGC.*GTAA" HBBmultifasta.txt
0
```

#ENZIMAS DE RESTRICCIÓN, SECUENCIAS GGATG, GGTGA y ACGT - SITIOS DE UNIÓN

```
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -E -o -n "GGATG|GGTGA|ACGT" HBBmultifasta.txt
3:GGTGA
3:ACGT
3:GGATG
3:GGTGA
5:GGTGA
6:ACGT
7:ACGT
13:GGTGA
14:GGATG
14:GGTGA
15:GGTGA
17:ACGT
23:GGTGA
24:GGATG
24:GGTGA
25:GGTGA
27:ACGT
36:GGTGA
36:ACGT
36:GGTGA
39:ACGT
40:ACGT
47:GGATG
51:ACGT
51:ACGT
51:GGTGA
52:GGTGA
59:GGTGA
59:ACGT
59:GGATG
61:GGTGA
63:ACGT
71:GGTGA
71:ACGT
71:GGATG
71:GGTGA
73:GGTGA
75:ACGT
75:ACGT
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -c "GGATG" HBBmultifasta.txt
6
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -c "GGTGA" HBBmultifasta.txt
15
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -c "ACGT" HBBmultifasta.txt
13
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -n "GGATG" HBBmultifasta.txt
3:GGAGAAGTCTGCCGTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGGAGCCCTGGGC
14:GGATCAAGTTGGTGGTGGAGCCCTGGGCAAGGTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTGA
24:CGTGGATCAAGTTGGTGGTGGAGCCCTGGGCAAGGTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTT
47:ATGGTGCACTGACTGAGGAGAAGTCTGCTTACTGCTGCTCCGCTGTGGGCAAGGTGGATGGAACAAG
59:ACCTGACTCTGAGGAGAAGTCTGCTTACTGCTGCTGGGCAAGGTGAACGTGGATCAAGTTGGTGG
71:TGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGCAAGGTGAACGTGGATCAAGTTGGTGGTGA
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -o -n "GGATG" HBBmultifasta.txt
3:GGATG
14:GGATG
24:GGATG
47:GGATG
59:GGATG
71:GGATG
(base) ubuntu@ubuntu-VirtualBox:~/media/sf_Compartida_VM/TrabajoBASH$ grep -n "GGTGA" HBBmultifasta.txt
3:GGAGAAGTCTGCCGTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGGAGCCCTGGGC
5:CTGTTATGGCAACCCCTAAGGTGAAGGCTCATGGCAAGAAAGTCTCGGTGCTTTAGTGATGGCTGGC
13:ACAGACACCCATGGTGCATCTGACTCTGAGGAGAAGACTGCCGTTACCACTCTGTGGGCAAGGTGAACG
14:GGATCAAGTTGGTGGTGGAGCCCTGGGCAAGGTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTTGA
15:TTCTCTTTGGGATCTGTCTCTCTGATGCTGTTATGGCAACCCCTAAGGTGAAGGCTCATGGCAAGAA
23:AAACACACACCATGCTGCACCTGACTCTGAGGAGAAGTCTGCTGTTACTGCCCTGTGGGCAAGGTGA
24:CGTGGATGAAGTTGGTGGTGGAGCCCTGGGCAAGGTGCTGGTGGTCTACCTTGGACCCAGAGGTTCTTT
25:GAGTCTCTTGGGATCTGTCCACTCTGATGCTGTTATGGCAACCCCTAAGGTGAAGGCTCATGGCAAGA
36:CCCTGTGGGCAAGGTGAACGTGGAGCAAGTTGGTGGTGAAGCCCTGGGCAAGGTGCTGGTGGTTACCC
51:GTGCAAGCTGCACTGGATCCTGAAATTTCAAGCTCTGGTGAAGTACTGGTGAATGTGCTGGCCCA
52:CCACTTTGGAAAGGAATTCCTGTGCCGTTGAGGCTGCCTATCAGAAAGTGGTGAAGTGGTGGCTAAT
59:ACCTGACTCTGAGGAGAAGTCTGCTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGG
61:TCCACTCTGATGCTGTTATGGGCAACCCCTAAGGTGAAGGCTCATGGCAAGAAAGTCTCGGTGCTTTA
71:TGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGA
73:ACTCCTGATGCTGTTATGGGCAACCCCTAAGGTGAAGGCTCATGGCAAGAAAGTCTCGGTGCTTTAGT
```

```
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ grep -o -n "GGTGA" HBBmultifasta.txt
3:GGTGA
3:GGTGA
5:GGTGA
13:GGTGA
14:GGTGA
15:GGTGA
23:GGTGA
24:GGTGA
25:GGTGA
36:GGTGA
36:GGTGA
51:GGTGA
52:GGTGA
59:GGTGA
61:GGTGA
71:GGTGA
71:GGTGA
73:GGTGA
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ grep -n "ACGT" HBBmultifasta.txt
3:GGAGAACTCTGCCGTTACTGCCCTCTGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAAGCCCTGGCC
6:TCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACCTGTGACAAGCTGCACGTGGAT
7:CTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCATCACTTTGGCAAAGAATTCA
17:AGCTGCACCTGTGACAAGCTGCATGTGGATCTGAGAAGTTCAAGCTCCTGGGCAACGTGCTGGTGTGT
27:TGAGCTGCACCTGTGACAAGCTGCACGTGGATCTGAGAAGTTCAAGCTCCTGGGCAAGTGTGGTGTGT
36:CCCTGTGGGGCAAGGTGAACGTGGAGAACTTGGTGGTGAAGCCCTGGGCAAGCTGCTGGTGGTTACCC
39:GCACCTTTGCCAGCTGAGTGAGCTGCACCTGTGACAAGCTGCACGTGGATCTGAGAATTTCAAGCTCCT
40:GGGCAACGTGCTGGTGTGTGTGGCTCATCACTTGGGCAAGAAATTCACCCGCAAGTGCAAGCTGCC
51:GTGACAAGCTGCACGTGGATCTGAAAATTTCAAGCTCCTGGGTAACGTACTGGTGATTGTCTGGCCCA
59:ACCTGACTCTGAGGAGAAGTCTGCTGTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGG
63:GCTGCACGTGGATCTGAGAAGTTCAAGCTCCTGGGCAAGTGTCTGGTCTGTGTCTGGCCCATCACTTT
71:TGACTCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGA
75:GCACGTGGATCTGAGAAGTTCAAGCTCCTGGGCAACGTGCTGGTCTGTGTCTGGCCCATCACTTTGGC
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ grep -o -n "ACGT" HBBmultifasta.txt
3:ACGT
6:ACGT
7:ACGT
17:ACGT
27:ACGT
36:ACGT
39:ACGT
40:ACGT
51:ACGT
51:ACGT
59:ACGT
63:ACGT
71:ACGT
75:ACGT
```

#ARCHIVO PDB - CADENAS DE LA PROTEÍNA

#ELIMINA TODAS LAS LÍNEAS QUE CONTENGAN INFORMACIÓN SOBRE PUENTES DE HIDRÓGENO + CREA NUEVO ARCHIVO

#CAMBIA PHE por ALA + NUEVO ARCHIVO

```
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ grep -c "TER" HBBprot.txt
0
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ grep -E -n -c "HETATM" HBBprot.txt
0
(base) ubuntu@ubuntu:~/VirtualBox:/media/sf_Compartida_VM/TrabajoBASH$ sed "s/F/A/g" HBBprot.txt>Cambiados.txt
```