|  |  |
| --- | --- |
| **Crèdit / Matèria: M4 Llenguatge de marques i sistemes de gestió de la informació** | **Data: Abril 2016** |
| **UF3: Sistemes de gestió empresarial** | **Curs i Grup: DAW-BIO** |
| **Nota**: |

**Pràctica 3 Perl: Fitxers**

Tenim en un fitxer de text les sequuències d'ADN dels diferents homínids (Les dades són inventades i el format és una simplificació del format FASTA):

adn.txt

> Homo sapiens

agatggcggcgctgaggggtcttgggggctctaggccggccacctactgg

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

> Homo erectus

agcggaagtagttgtgggcgcctttgcaaccgccggggtcttgggggctc

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

> Homo habilis

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

agatggcggcgctgaggggtcttgggggctctaggccggccacctactgg

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct

La línea que comença amb el caràcter > ens proporciona el nom de la seqüència.

Volem separar cada seqüència en un fitxer diferent que tindrà el nom de la seqüència i el seu contingut . El l'exemple el resultat ens donaria tres fitxers de sortida:

Homo sapiens.txt

agatggcggcgctgaggggtcttgggggctctaggccggccacctactgg

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

Homo erectus.txt

agcggaagtagttgtgggcgcctttgcaaccgccggggtcttgggggctc

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

Homo habilis.txt

gggaggcgtgactagaagcggaagtagttgtgggcgcctttgcaaccgcc

agatggcggcgctgaggggtcttgggggctctaggccggccacctactgg

tttgcagcggagacgacgcatggggcctgcgcaataggagtacgctgcct