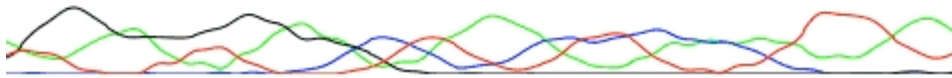


## Análise Interna de Sequências de DNA

### Depuração de Sequências

1 - Considere o seguinte fluorograma que acabou de receber:

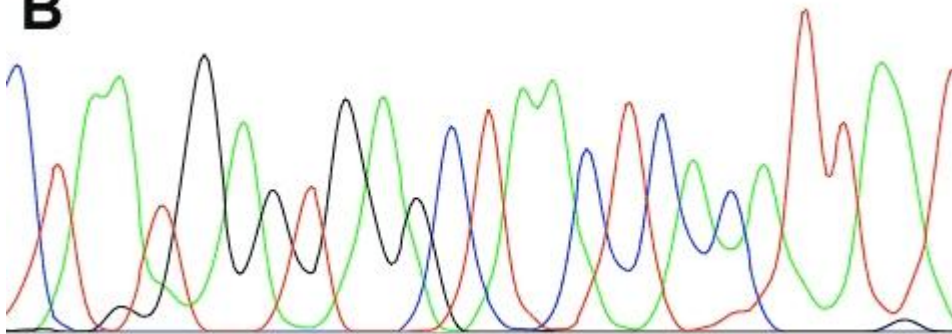
**A**



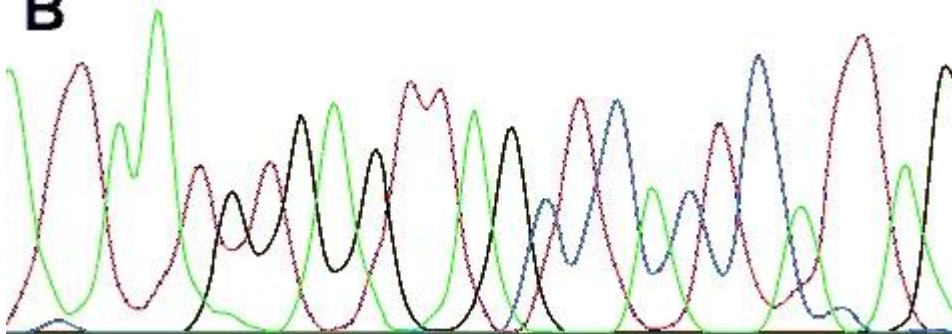
Sabendo que verde é A, vermelho T, azul C e negro G interprete-o em termos de sequência e guarde o resultado em formato FASTA. Se necessário recorra a códigos de ambiguidade da IUPAC.

2 - O fluorograma anterior corresponde à extremidade final de uma leitura longa. Para mitigar o baixo valor do quociente sinal/ruído é vulgar normalizar cada um dos sinais de acordo com a sua amplitude. O resultado é mais legível, embora possa esconder artefactos:

**B**



**B'**



Repita a interpretação, registre as sequências correspondentes em formato FASTA e determine qual a relação entre elas.

!

## Depuração de Sequências

```

Watson      -----
Crick      AGGAYTACCCGCTGAACTTAVVGCATATCAWTAVGCGGGGGAAAAGAACTAACAAGGAT
              190      200      210      220      230      240

              10      20      30      40      50
Watson      TTCCCTAGTAGCGGCGAGCGA-GSNGGAAGAGCTCAAATTTATAATCTGGGCACCTTCGG
              :  :::::::::::::::::::: :  :::::::::::::::::::: :  :::::
Crick      TCCCCTAGTWGCGGCGAGCGAAGCGGGAAGAGCTCAAATTTATAATCTGGCM--CTTCGG
              250      260      270      280      290

```

1 - Imagine que acabou de obter e comparar o texto correspondente às duas cadeias complementares de uma sequência de DNA. Considere o fragmento de comparação acima apresentado.

2 - Para finalizar edite a sequência que alinhou e resolva as ambiguidades dando prioridade aos resíduos compatíveis com a ambiguidade na cadeia oposta (na realidade teria de considerar os dados de fluorescência). Como aproximação ignore G/C emparelhado com lacuna e no caso oposto considere o A ou o T presente.

3 - Guarde o resultado num ficheiro segundo o formato FASTA.

## Armazenamento

```

>A1E2E6|A1E2E6_9TELE| Growth hormone;
MARVLVLLSLAMASLYVSQGSTLESQRLFNNAVARVQHLHQLAAKMMDDFDEALLPEERK
QLGKIFPLSFCNSDSIEAPAGKDETQKSSVLKLLHTSYRLIDSWDFPSRTLGNPNHISDK
LADLKIGISVLIERCVDGQTTLDEIDSLAPPFEDFYQTPSEGNLSKHFRLLSCFKKDMHK
VETYLSVAKCRRSLDSNCTL
>B7TK82|B7TK82_9TELE| Growth hormone;
LLSVVLVSVFVNNGTASENQRLFNNAVIRVQHLHQLAAKIINDFEDSLLPEERRQLSKIF
PLSFCNSDSIEAPTGKDETQKSSVLKLLRISFRLIESWEYPSQTLSGTISNSLTIGNPNQ
ITEKLADLVKGISVLINGCLDGQPNMDDNDSLPLPFEDFYLTGENSELRESFRLACFKK
DMHKVETYLVRVANCRRSLDS
>B7TKA6|B7TKA6_9TELE| Growth hormone;
VLVSLVQNQGTASENQRLFNNAVIRVQHLHQLAAKMINDFEDNLLPEERRQLSKIFPLSF
CNSDSIEAPTGKDETQKSSMLKLLRISFRLIESWEFPSQTLSGAISNSLTIVGNPNQITEK
LADLVKGISVLKGLDGQPNMDENDSLPLPFEDFYLTMGESSLRESFRLACFKKDMHK
VETYLVRVANCRRSLDSNGTL
>D3Y3L8|D3Y3L8_9TELE| Growth hormone;
MARVLVLLSVVLVSLFFSQGATFENQRLFNNAVIRVQHLHQLAAKMMDDFEEALLPDERK
QLSKIFPLSFCNSDSIEAPAGKDETQKSSVLKLLHTSYRLIESWEFPSKNLGNPNHISEK
LADLMGIGVLIIEGCLDGQTSMDENDALAPPFEDFYQTLSEGNLRKSFRLSCFKKDMHK

```

```

VETYLTVAKCRRSLDSNCTL
>F1B1D9|F1B1D9_MORAM| Growth hormone;
VLLLSVLSSLGVSSQPITEGQRLFSIAVERVHNLHLLAQRLTFEFESSLQTEEQRLNKIF
LQDFCNSDYIISPIDKHETQRSSVLKLLSISYRLIESWEFPPSRSLSVGPAARNQISPKLS
ELKTGILLIGANQDGAEMFPDSSLQAPYGNYYQSLGADESLRRTYELLACFKKDMHK
VETYLTVAKCRRSLDSNCTL
>H6T596|H6T596_9PERC| Growth hormone;
MDKLVLLLSMLCMSVSTQPVIDQRRFSIAVSKVQHLHLLAQKYFSDFENSLQTEDQQRQFN
KFHQDFCNPNDNIIIPDNKHETQRSSVLNLLLISKRLVDSREFSIHFITWNLFPNQLVHR
LSDLKKGIGMLIEASDGAEMSDSITLPVTPFGNFYENLGGNESQKRNYELLACFKKDMHK
VETYLTVAKCRRSLDSNCTL
>Q8AV82|Q8AV82_SILAS| Growth hormone;
MARVWVLLSVVLASLFFSQGATFENQRLFNNAVIRVQHLHQLAAKMMGDFEEALLPEERK
QLSKIFPLSFCNSDSIEAPSGKDETQKSSVLKLLHTSYRLIESWEFPPSKNLGNPNHISEK
LADLKMIGIGVLIIEGCMGQTSLEDENDSLAPPFEDFYQTLTEGNLRKSFRLLSCFKKDMHK
VETYLTVAKCRRSLDSNCTL
>Q8QGJ1|Q8QGJ1_HETFO| Growth hormone protein;
MARVLVLLSVVVASLFFSQGATFETQRLFNNAVIRVQHLHQLAAKMMDDFEEALLPEERK
QLSKIFPLSFCNSDSIEAPAGKDETQKSSVLKLLHTSYRLIESWEFPPSKNLGNPNHISEK
LADLKMIGIGVLIIEGCVDGQTSLEDENDAFAPPFEDFYQTLSEGNLRKSFRLLSCFKKDMHK
VETYLTVAKCRRSLDSNCTL

```

1 - Transforme as sequências numa tabela com quatro colunas (p.e usando um editor de texto), use ';' como separador e guarde o resultado.

2 – Instale e execute o Sqlliteman Portable (<http://portableapps.com/apps/development/sqlliteman-portable>).

3 – Crie uma nova base de dados (File|New).

4 – Crie uma nova tabela com colunas destinadas a receber as suas sequências (p.e.

```

CREATE TABLE "sequeces" (
    "ac" TEXT NOT NULL,
    "id" TEXT,
    "description" TEXT,
    "sequence" TEXT NOT NULL,
)

```

5 – Importe as suas sequências selecionando a tabela no esquema e importando a tabela a partir do ficheiro criado em 1 (p.e. Database|Import Table Data).

6 – Verifique no painel da direita que a tabela tem agora as suas sequências.

7 – Recupere as sequências cujo 'ac' inclua '6' (p.e. Database|Build query; selecione 'ac', 'sequence').