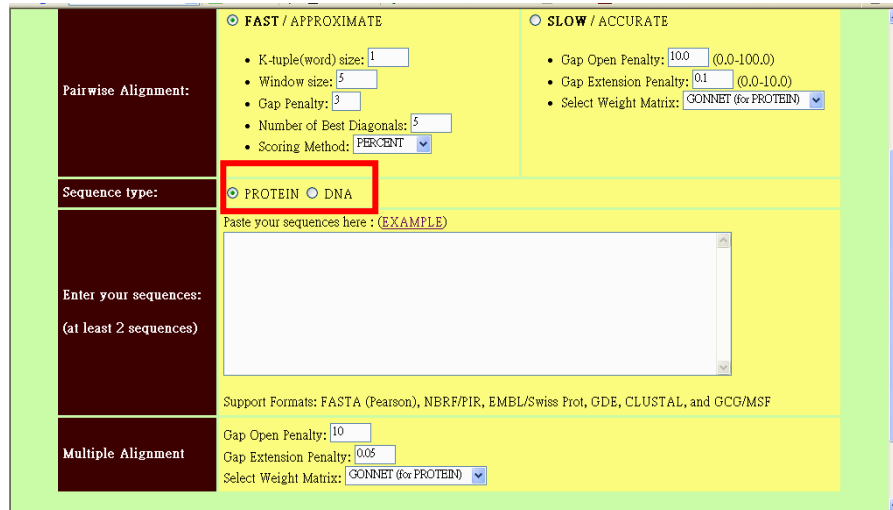


ClustalW

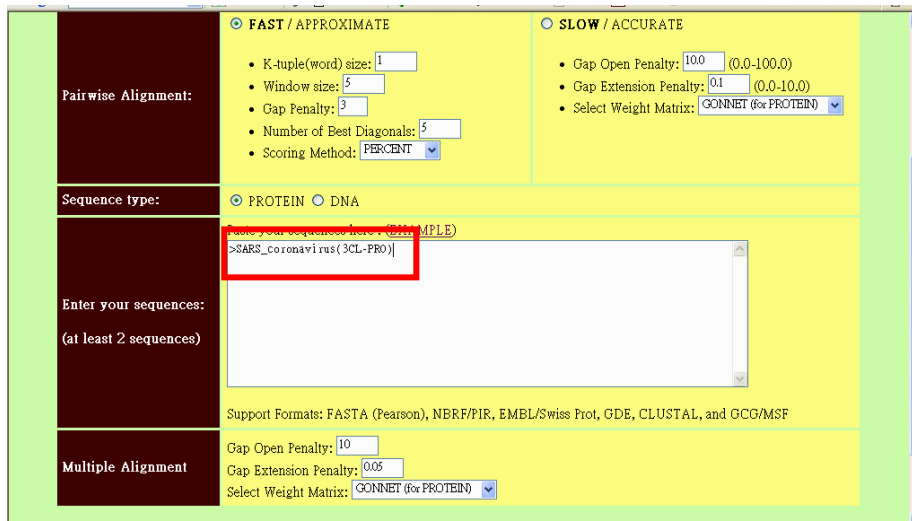
1. 序列比對分為四步驟：

- (1) 選取比對序列的特性，在此有分為”FAST”和”SLOW”的計算方式，”SLOW”是計算時間較久但結果相對地也較正確，”FAST”是計算時間較快但結果相對地也較不正確。另外，在輸入序列前，先指定是蛋白質的序列或 DNA 的序列。



The screenshot shows the ClustalW web interface. The 'Sequence type' section has two radio buttons: 'PROTEIN' (selected) and 'DNA'. The 'FAST / APPROXIMATE' section is active, showing parameters like K-tuple(word) size: 1, Window size: 5, Gap Penalty: 3, Number of Best Diagonals: 5, and Scoring Method: PERCENT. The 'SLOW / ACCURATE' section is inactive, showing parameters like Gap Open Penalty: 100, Gap Extension Penalty: 0.1, and Select Weight Matrix: GONNET (for PROTEIN). The 'Enter your sequences' text area is empty. The 'Multiple Alignment' section shows parameters like Gap Open Penalty: 10, Gap Extension Penalty: 0.05, and Select Weight Matrix: GONNET (for PROTEIN).

- (2) 輸入序列名稱及 PDB code (若不知 PDB code 可省略)，也就是輸入：”> + 名稱 + PDB code”，例如：”> SARS_coronavirus(3CL-PRO)”，切記在序列名稱前加大於”>”符號！



The screenshot shows the ClustalW web interface. The 'Sequence type' section has 'PROTEIN' selected. The 'Enter your sequences' text area contains the input '>SARS_coronavirus(3CL-PRO)'. The 'FAST / APPROXIMATE' section is active, showing parameters like K-tuple(word) size: 1, Window size: 5, Gap Penalty: 3, Number of Best Diagonals: 5, and Scoring Method: PERCENT. The 'SLOW / ACCURATE' section is inactive, showing parameters like Gap Open Penalty: 100, Gap Extension Penalty: 0.1, and Select Weight Matrix: GONNET (for PROTEIN). The 'Multiple Alignment' section shows parameters like Gap Open Penalty: 10, Gap Extension Penalty: 0.05, and Select Weight Matrix: GONNET (for PROTEIN).

(3) copy/paste : 直接貼上序列

The screenshot shows the 'Pairwise Alignment' section of a software interface. It has two columns of settings: 'FAST / APPROXIMATE' and 'SLOW / ACCURATE'. The 'FAST' column includes: K-tuple(word) size: 1, Window size: 5, Gap Penalty: 3, Number of Best Diagonals: 5, and Scoring Method: PERCENT. The 'SLOW' column includes: Gap Open Penalty: 100 (0.0-100.0), Gap Extension Penalty: 0.1 (0.0-10.0), and Select Weight Matrix: GONNET (for PROTEIN). Below these is the 'Sequence type' section with 'PROTEIN' selected. The 'Enter your sequences' section has a text box containing a protein sequence:

```
>SARS_coronavirus(3CL-PRO)
>SARS_coronavirus(3CL-PRO)
GFRKMAFFSGKVEGCMQVTCGTTLNLGLMLDDTVYCPRHVICTAEDMLNPVYEDLLIRKSNHSFLVQA
NFVQLRVIQHSNMQNCLRLKVDTSNPKTKPKYKFRVIRIQGQTFSVLACYNQSPSGVYQCAMPNHTIKGSF
LN3SCGSGVFNIDYDCVFSFCYHMHMELPT3VHAGTDLEGRFVGFVDRQTAQAAGTDTTTLAVLAWLYA
VINGDRWFLNRFITTLNDFNLVAMKYNVEPLTQDHVDILGPLSAQT3IAVLDMCAALKELLQNMGRGT
LGSTILEDEFTPFDDVVRQCSGVTFQ
```

 The first two lines are highlighted in red. Below the text box is the 'Multiple Alignment' section with settings: Gap Open Penalty: 10, Gap Extension Penalty: 0.05, and Select Weight Matrix: GONNET (for PROTEIN). Support Formats listed are FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF.

(4) 依此方法再將其他需進行比對的序列輸入，可輸入多條的蛋白質序列。

This screenshot shows the same software interface as above, but with multiple sequences entered in the 'Enter your sequences' text box. The sequences are:

```
>SARS_coronavirus(3CL-PRO)
>SARS_coronavirus(3CL-PRO)
GFRKMAFFSGKVEGCMQVTCGTTLNLGLMLDDTVYCPRHVICTAEDMLNPVYEDLLIRKSNHSFLVQA
NFVQLRVIQHSNMQNCLRLKVDTSNPKTKPKYKFRVIRIQGQTFSVLACYNQSPSGVYQCAMPNHTIKGSF
LN3SCGSGVFNIDYDCVFSFCYHMHMELPT3VHAGTDLEGRFVGFVDRQTAQAAGTDTTTLAVLAWLYA
VINGDRWFLNRFITTLNDFNLVAMKYNVEPLTQDHVDILGPLSAQT3IAVLDMCAALKELLQNMGRGT
LGSTILEDEFTPFDDVVRQCSGVTFQ
>Murine_hepatitis_virus(3CL-PRO)
>Murine_hepatitis_virus(3CL-PRO)
SGIVKIVSPTSKEVPCIVSVTYGNMILNGLMLDDKVVYCPRHVICSSADMTDPPVPLLCRVTSDFCVMS
GRMSLTVMISVQMGCCQLMLVTLQNPNTPKYFVGVVKGPEFTVLAAYNGRPGGAFHVLRSHTIKGSF
LCGSCGSGVYVLTGDSVRFVYMHQLELSTGCHTGTDFSGNFVGPYRDAQVWQLPVQVYVTVNVAWLYA
```

 The second sequence is highlighted in red. The rest of the interface, including settings and support formats, remains the same as in the previous screenshot.

