

DNAMAN

Sequence Analysis Software

*Sequence
Alignment and
Phylogenetics*

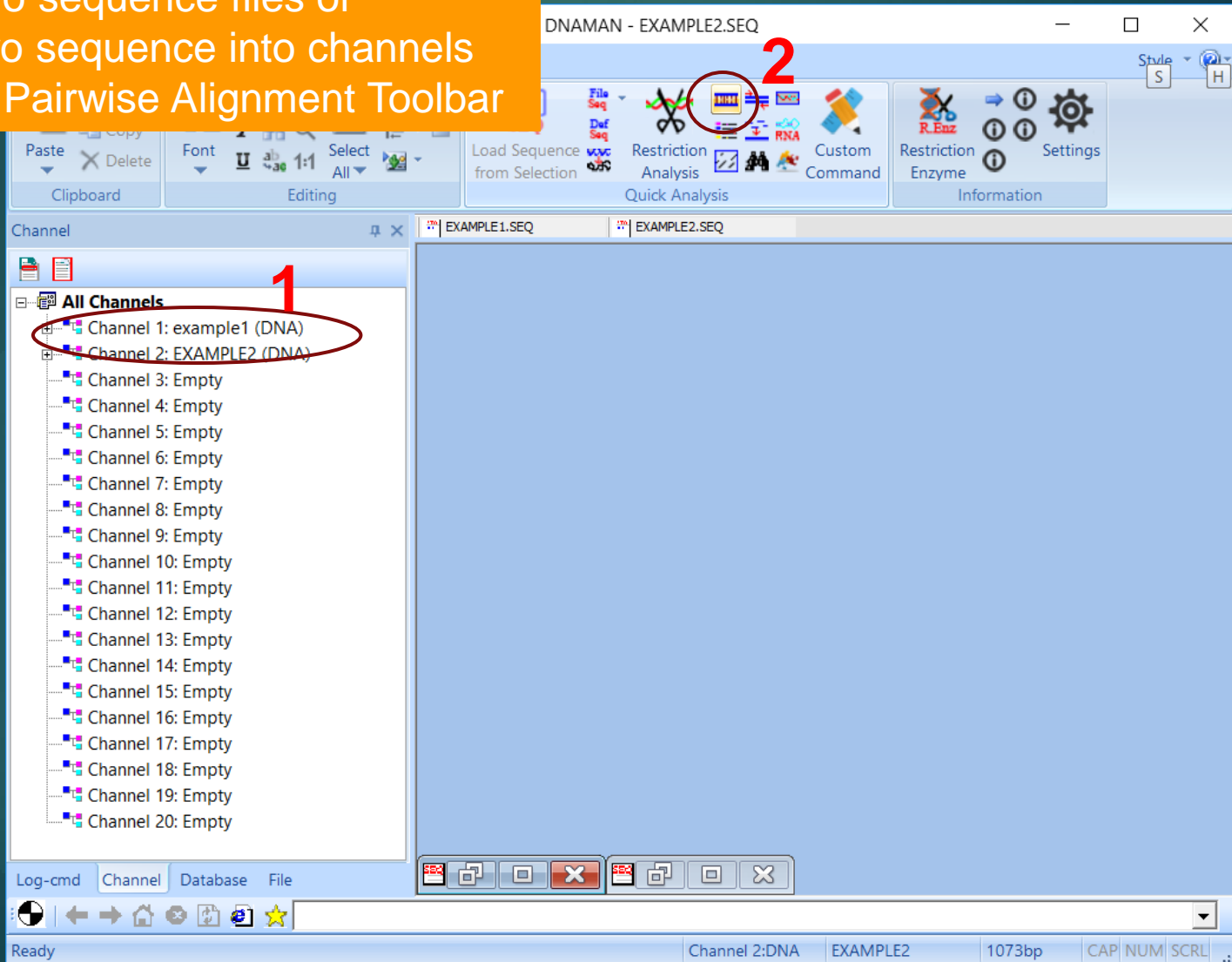
BIOINFORMATICS PLATFORM

Sequence Alignment and Phylogenetics

- ▶ Pairwise Sequence Alignment
- ▶ Multiple Sequence Alignment
- ▶ Phylogenetic Analysis

Pairwise Alignment

1. Open two sequence files or load two sequence into channels
2. Choose Pairwise Alignment Toolbar



Pairwise Alignment

1. Select two sequences from channels
2. Select type of alignment
3. If DNA seq properties unknown, check “Best alignment of dsDNA”
4. For Protein Seq alignment, choose a score matrix
5. Choose one of 4 alignment method
6. Choose parameters for alignment
7. Select sequence alignment display options
8. Click OK to start alignment

The screenshot shows the 'Pairwise Alignment' dialog box with the following settings and annotations:

- 1**: Sequence 1 dropdown menu set to 'CH. 2:EXAMPLE2'.
- 2**: Radio button for 'DNA' selected.
- 3**: Check box for 'Best alignment of dsDNA' checked.
- 4**: Score Matrix dropdown menu set to 'BLOSUM'.
- 5**: Radio button for 'Quick Alignment' selected.
- 6**: 'Alignment Method' section containing:
 - Gap open: 2
 - Gap extension: 1
 - K-tuple: 2
 - Window: 4
 - Diagnols: 4
 - Myers&Miller (global) and Needleman&Wunsch (global) are unselected.
 - Smith&Waterman (local) is unselected.
- 7**: 'Show Results' section containing:
 - 'Show with symbol' is checked.
 - 'Identity' radio button is selected.
 - 'Mismatch' radio button is unselected.
 - 'Show flanking sequence length <' is checked, with a value of 60.

Buttons for 'OK' and 'Cancel' are at the bottom.

Pairwise Alignment

Result of Pairwise Alignment

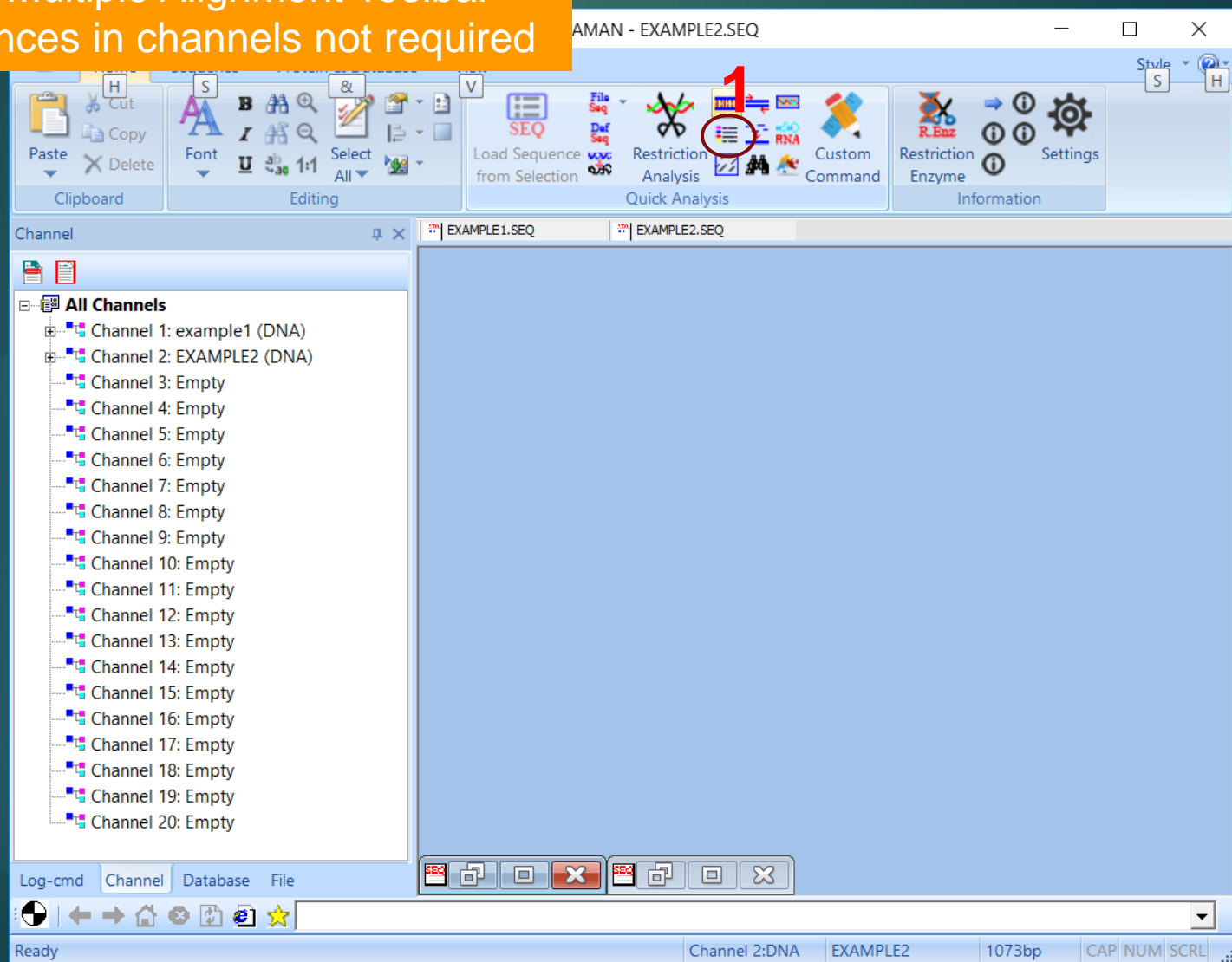
The screenshot displays the DNAMAN software interface. The main window shows a global alignment of two DNA sequences: EXAMPLE2 (top) and example1 (bottom). The alignment is performed using the Needleman-Wunsch algorithm. The top sequence is EXAMPLE2:example1 with an identity of 79.89% (695/870) and a gap of 16.83% (176/1046). The bottom sequence is example1. The alignment shows the two sequences with vertical bars indicating matches and gaps. The sequences are as follows:

```
EXAMPLE2:example1 identity= 79.89%(695/870) gap=16.83%(176/1046)
1 .....CACCTTCCTTGACGAGGGCTTTACTGCCAAGGATATCCTCGACCAAAAAATA
1 TTTGACTGCCACTTCCTCGATGAAGTTTTACTGCCAAGGACATTCTGGACCAGAAAATT
53 AACGAAGTGTCACTTCTGATGATAAAGATGCCTTCTATGTTGCTGACCTCGGGGATATT
61 AATGAAGTTTCTTCTTCTGATGATAAAGGATGCCTTCTATGTGGCAGACCTGGGAGACATT
113 GTAAAGAGACACATGCGGTGGCATAAAGCCCTTCCTCGAGTAACCCCTTCTACGCTGTC
121 CTAAAGAAACATCTGAGGTGGTTAAAAGCTCTCCCTCGTGTCAACCCCTTTTATGCAGTC
173 AAATGGTAATCGACAGTCAGCTTTCACGCTTGCAGTTAATATCATTGCCAAGAAAATTGT
181 AAAT.GTAA.....TGATAGCAAAGCCAA.....
233 ATTAAGGAACAGACGGGCTCTGATGACGAAGATGAGTCGAGTGAGCAGACCTTTATGTA
204 .....GACGAAGACGAG.AGAATGACAAAACACTCATGTA
293 TTATGTGAATGATGGCGTCTATGGATCAITTAATGCACTACTCTATGACCACGCACATGT
238 TTACGGGAATGATGGTGTCTATGGATCGTTCAATTGCATCTTGATGATCATGCACATGT
353 AAAGCCCTTCTGCAAAAGAGACCTAAACCAGATGAGAAGTATTATTCAITCCAGCATATG
298 TAAACCAGTTCTGCAAAAGGGCTTAAACCAGATGACGGCTGCTACTCCTGCAGCATATG
```

The interface includes a menu bar (Home, Sequence, Protein & Database, View), a toolbar with various editing and analysis tools, and a channel list on the left. The status bar at the bottom shows 'Ready', 'Channel 2:DNA', 'EXAMPLE2', '1073bp', and 'CAP NUM| SCRL'.

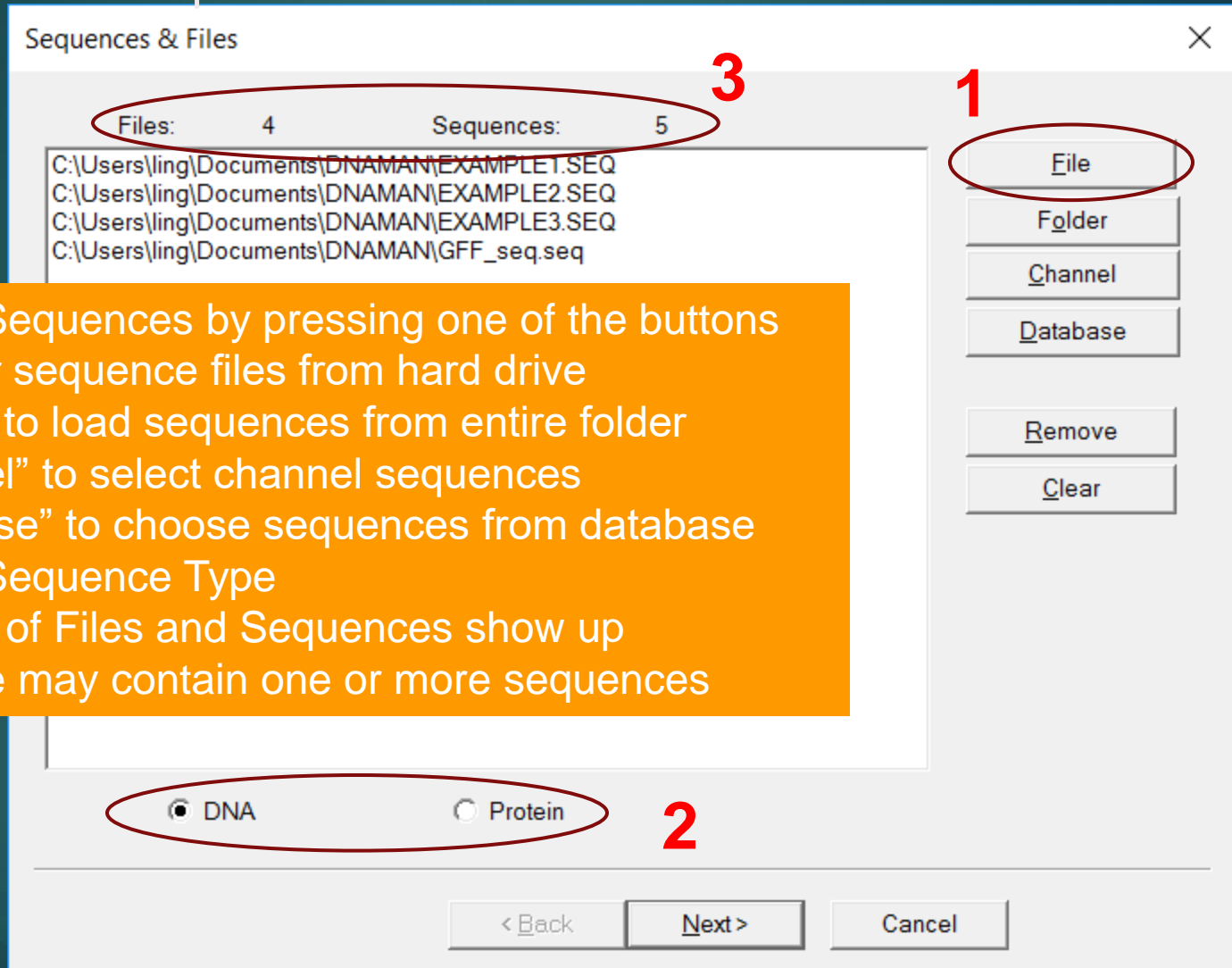
Multiple Alignment

1. Choose Multiple Alignment Toolbar
Sequences in channels not required



Multiple Alignment

- Select Sequences



1. Choose Sequences by pressing one of the buttons
"File" for sequence files from hard drive
"Folder" to load sequences from entire folder
"Channel" to select channel sequences
"Database" to choose sequences from database
2. Choose Sequence Type
3. Numbers of Files and Sequences show up
Each file may contain one or more sequences

Multiple Alignment

- Alignment Methods

1. Choose Alignment Methods from 3 Optimal and 1 Fast alignment

“Full Alignment” to fully align all sequences

“Profile Alignment” to align two sets of aligned sequences

“New Sequences on Profile” to align more sequences to the first sequence set

“Fast Alignment” to align all sequences using fast alignment algorithm

Method

Optimal Alignment

Full Alignment **1**

Profile Alignment

New Sequences on Profile

Fast Alignment

Output order

Input **2** Aligned

Translation from DNA

No **3** Frame 2

Frame 1 Frame 3

Try both strands

Run in background

Show progress

Multiple Alignment

- Multiple Alignment Parameters

1. To use default parameters instead of user-defined ones, press Default Parameters button
2. Change Gap Open and Extension penalties if needed
3. Sequences with poor homologous score can be delayed for alignment
4. Protein specific weight matrix can be selected
5. Negative matrix and penalties for gaps at ends can be used
6. More protein gap penalties can be used

The screenshot shows the 'Multiple Alignment' dialog box with the following parameters and annotations:

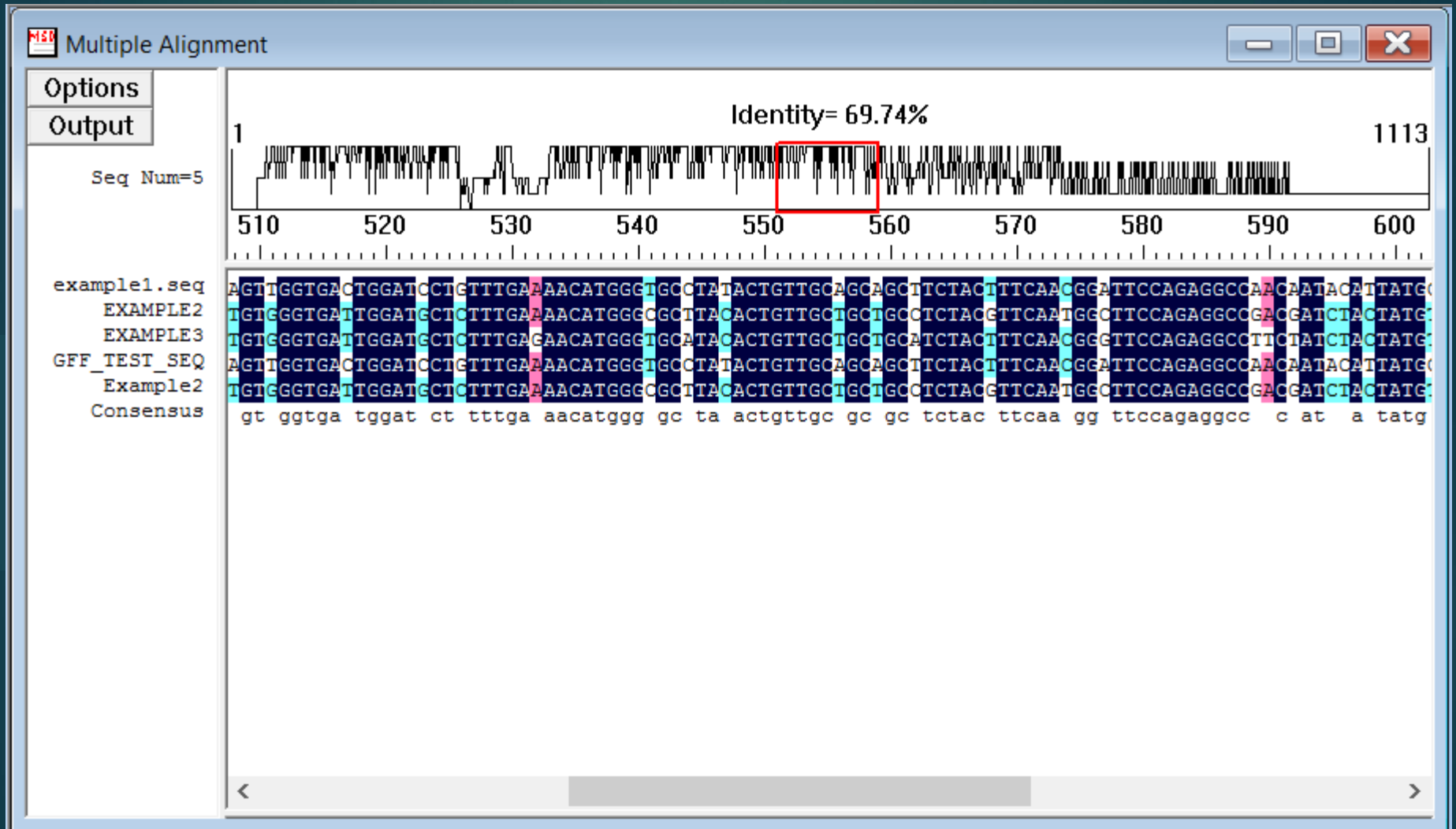
- 2**: Gap Open Penalty: 10
- 2**: Gap Extension Penalty: 5
- 5**: Negative matrix
- 3**: Delay Divergent Seqs%: 30
- 4**: Protein Weight Matrix: GONNET
- 4**: Penalty End Gaps
- 1**: Default Parameters button
- 6**: Protein Gap Parameters section containing:
 - Penalty Hydrophilic Residues: GPSNDQEKR
 - Gap Separation Distance: 4
 - Penalty Residues by Pascarella Probabilities

Buttons at the bottom: < Back, Finish, Cancel

Reference:
Thompson et al (1994) *Nucleic Acids Res.* 22:4673

Multiple Alignment

- Multiple Alignment Result



Phylogenetic Analysis

The screenshot shows the 'Multiple Alignment' software interface. The 'Output' menu is open, and the 'Tree' option is selected. The 'Phylogenetic Tree' sub-option is also highlighted. The background shows a sequence alignment with an identity of 69.74% and a scale bar from 0 to 80. The sequence alignment is as follows:

```
example1.s  
EXAMP1  
EXAMP1  
BFF_TEST_5  
Examp:  
Consens:
```

Identity= 69.74%

0 80

ACTTCCTGATGAGGCTTTTACGCCAAGGACATTCCTGACCAGAAAATTAATG
CCTTCCTGACGAGGGCTTTTACGCCAAGGATATCCTGACCAGAAAATTAATG
ATTCCTGATGAGGCTTTTACGCCAAGGACATTCCTGACCAGAAAATTAATG
ACTTCCTGATGAGGCTTTTACGCCAAGGACATTCCTGACCAGAAAATTAATG
CCTTCCTGACGAGGGCTTTTACGCCAAGGATATCCTGACCAGAAAATTAATG

tcct ga ga gg tttac gccaagga at ct gacca aaaat aa g

1. Click Output button
2. Select Tree | Phylogenetic Tree menu

Phylogenetic Analysis

- Tree Options

Options

Distance Methods

Observed Divergency

Kimura

Jukes & Cantor (DNA only)

Poisson Correction (Protein only)

Maximum Likelihood

Toss gaps

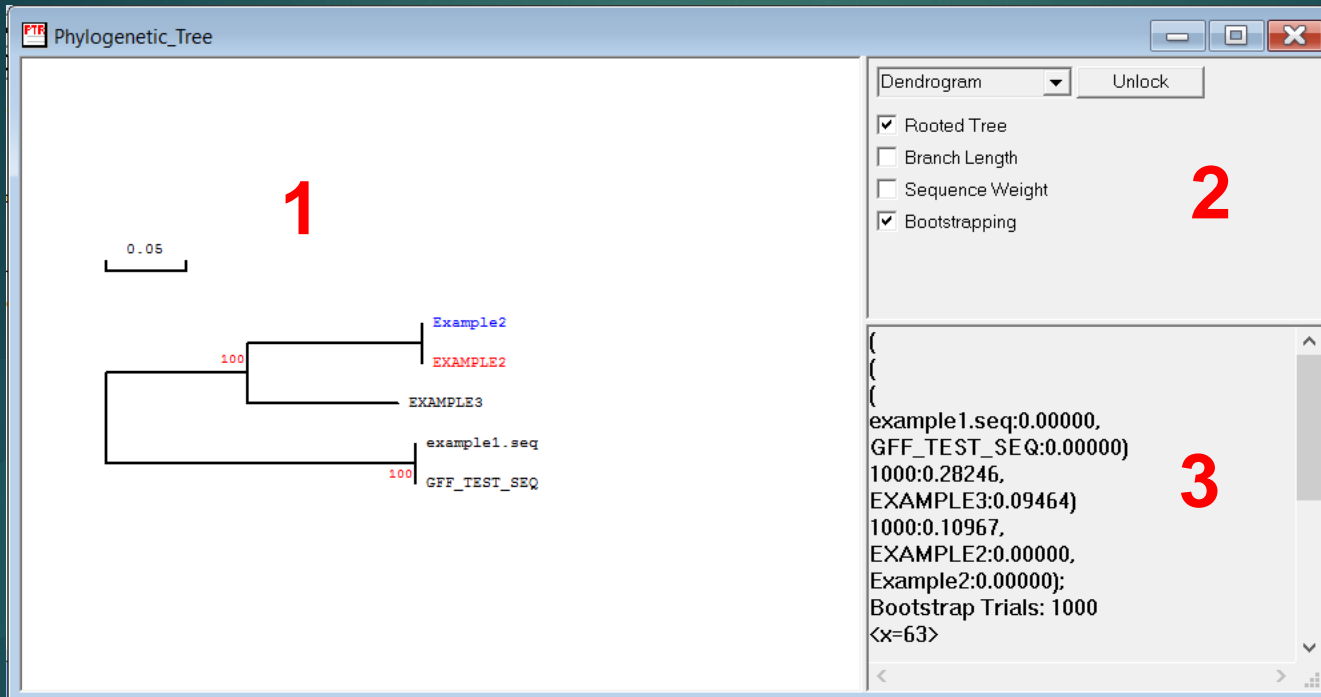
Bootstrap trials (0-1000000)

< Back Finish Cancel

1. Choose Kimura method (optional)
2. Enter bootstrap trial number (optional)

Phylogenetic Analysis

- Tree View



1. Tree graph window: Double-click to open Tree Options dialog box
2. Graph control window
3. Tree description text window

Phylogenetic Analysis

- Tree View Options

Options

Tree Type

Dendrogram Circular Spindle

Show

Rooted Tree Scale

Branch length

Sequence weight

bootstrapping value \geq %

Font

Replace _ by space in seq name

Alter italic font for strain name

Colour Branch Level

Graph Position: X

Graph Size: Width

Root length

Y

Branch space

Phylogenetic Analysis

- Tree Type

