Align2

Why was written and what does "Align2"

The program 'Align2' has been developped to facilitate the comparison of DNA banding patterns after their transformation in lists of band size values as it is common when using automatic reading device such as sequencing apparatus. The program was used to find DNA bands reliably obtained with the same bacterial genome in different AFLP analyses (Cf. Mougel *et al*. 2001, 2002)

The program ranks DNA band sizes on the same scale, gives the same rank to DNA bands having the same size (± 1 bp) or leaves blanks if DNA bands are found in one pattern and not in the other pattern.

Input data are two files made of the respective lists of band size values expressed in base pairs. Band sizes can be real values as it is commonly obtained with outputs of automatic reading devices.

The output data are average values of aligned size data.

Format of input and output files

Input files

Two input files are required, one for each DNA pattern (ex: Profile1 and Profile2)
Input files contain the list of DNA band sizes (in bp) bands separated by "return"
Input files must be saved in "text only" format.

Output file

The name of the output file is given by the program by joining the names of the two input files:
                        (ex: Profile1+Profile2_align)
The output file contains three columns and can be opened in a spreadsheet.
First and second columns contain ranked and aligned values of the first and second input files, respectively. The third column contains the average size values of DNA bands common to both patterns.

Running Align2

A dialog box is open asking for file 1 and file 2. Those files must be in the current directory, else the directory can be selected using "Preference" in the "Edit" menu.

Example

Input file 1:   ex: Profile1

```
12.8
13
15
100
110
115.25
156
158
1308.2
```

Input file 2:   ex: Profile2

```
12
12.8
14
15
25
35
100
115.25
155.55
157.2
1200
1100
1308.2
```

Output file: ex: Profile1+Profile2_align

```
12          12.8        12.4
12.8        0           0
13          12.8        12.9
0           14          0
15          15          15
0           25          0
0           35          0
100         100         100
110         0           0
115.25      115.25      115.25
156         155.55      155.775
158         157.2       157.6
0           1200        0
0           1100        0
1308.2      1308.2      1308.2
```

References

Mougel C., Teyssier S., d'Angelo C., Groud K., Neyra M., Sidi-Boumedine K., Cloeckaert A., Peloille M., Baucheron S., Chaslus-Dancla E., Jarraud S., Meugnier H., Forey F, Vandenesch F., Lina G., Etienne J, Thioulouse J., Manceau C., Robbe P., Nalin R., Briolay J., Nesme X. 2001. Experimental and theoretical evaluation of typing methods based upon random amplification of genomic restriction fragments (AFLP) for bacterial population genetics. Genet. Sel. Evol. 33 (Suppl. 1):S319-S338.

Mougel C., Thioulouse J., Perrière G., Nesme X. 2002. A mathematical method for determining genome divergence and species delineation using AFLP. Int. J. Syst. Evol. Microbiol. 52:573-586.