

Iterative Refinement Unit Tests

Hint: Many test values are taken from project Algorithms for Bioninformatics of Alexander Mattheis or the lectures.

Test 1 (used Feng-Doolittle implementation and Sum-of-Pairs implementation)

Input

Sequence a: ACGT
Sequence b: AT
Sequence c: GCT
Sequence d: GC

Gap opening: -1
Enlargement: -3
Match: 1 (and 0 for placeholder #)
Mismatch: 0

Output (Computation: Global Primary Library)

Feng-Doolittle Output

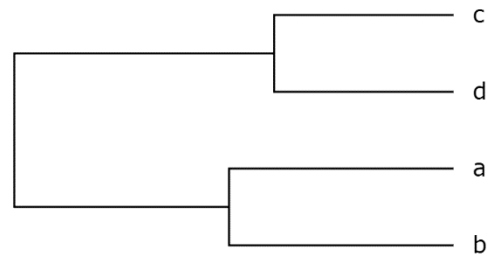
Alignment:

ACGT
A__T
G_CT
G_C_

Distance-Table:

D	a	b	c	d
a	0	1.8	1.7	2.9
b	1.8	0	2.2	7.6
c	1.7	2.2	0	1.5
d	2.9	7.6	1.5	0

Tree:



Round-Robin – Minimum Distance Pair Output

Post-Order-Traversal: badc

removed

sequence (Hint: is cleaned from gaps)

b: $\operatorname{argmin}_x (D_{b,x}) = a$ with $D_{b,a} = 1.8$

b~a: AT~ACGT

A__T
__

ACGT

Nothing changes!

a: $\operatorname{argmin}_x (D_{a,x}) = c$ with $D_{a,c} = 1.7$

a~c: ACGT~GCT

ACGT
|*_*
GC_T

remaining
alignment

ACGT
G_CT
G_C_

A__T
GCT

GC_

(gap-only columns removed)

after realignment:

ACGT
A__T
GC_T
GC__
Score: -26

Hint: This alignment is kept.

d: $\operatorname{argmin}_x(D_{d,x}) = c$ with $D_{d,c} = 1.5$

d~c: GC~GC#T

GC__
**
GC#T

ACGT
A__T
GC_T

Nothing changes!

c: $\operatorname{argmin}_x(D_{c,x}) = d$ with $D_{c,d} = 1.5$

c~d: GC#T~GC

GC#T
**
GC__

ACGT
A__T
GC__

Nothing changes!

Feng-Doolittle Alignment:

ACGT
A__T
G_CT
G_C_

Refined Realignment:

ACGT
A__T
GC_T
GC__

Round-Robin – One-Vs-All Output

(analogous - but it is searched for the alignment which produces best MSA-score)

Round-Robin – Pairwise Best Pairs Output

(analogous - but it is searched for the best and not the nearest alignment)

The other used procedures are already have been tested during T-Coffee and Feng-Doolittle.