

## Supplementary information

The `-leavegappyregion` option This option uses a position-specific gap cost  $G_1(i, x)$  for group-to-group alignment,

$$G_1(i, x) = S^{\text{op}}\{(1 - g_1^{\text{start}}(x)) + (1 - g_1^{\text{end}}(i))\} / 2, \quad (9)$$

where  $S^{\text{op}}$  is a normal gap cost for sequence-sequence alignment,  $g_1^{\text{start}}(x)$  is the frequency of gaps that start at position  $x$ , and  $g_1^{\text{end}}(i)$  is the frequency of gaps that end at position  $i$ . See Katoh *et al.* (2002) for details. This was the default gap cost of MAFFT till 2013 Oct.

In the current version (since v7.113, 2013 Oct), the default gap cost is:

$$G_1(i, x) = S^{\text{op}}\{(1 - g_1^{\text{start}}(x)) f(x) + (1 - g_1^{\text{end}}(i)) f(i)\} / 2, \quad (10)$$

where  $f(x)$  is the frequency of non-gap characters at position  $x$ .

Obviously, the gap cost of equation 9 is stronger than that of equation 10. By using the former, less gaps are inserted to gap-rich regions, which can be discarded manually or by a filtering program. Such an MSA is sometimes useful when one is interested only in mostly gap-free sites. For large input data consisting of thousands or more sequences, the current default (eq. 10) outperforms the previous default in sensitivity.

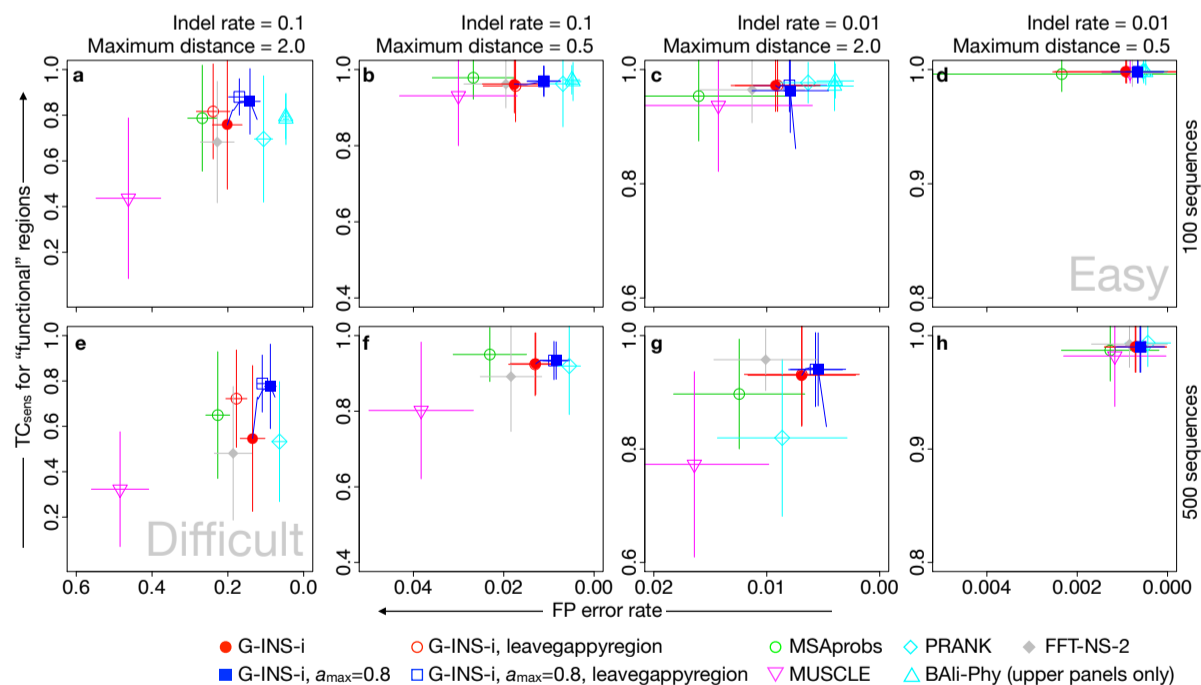
### Setting for INDEliBLE simulations

[TYPE] AMINOACID 2

[SETTINGS]

[output] FASTA

```
[randomseed] (integer)
[MODEL] modelname
[submodel] WAG
[rates] 0.0 1.0 0
[indelmodel] POW 1.7 500
[indelrate] 0.01 (or 0.1)
[MODEL] modelcons
[submodel] WAG
[rates] 0.0 1.0 0
[indelmodel] POW 1.7 500
[indelrate] 0.00
[TREE] treename
(randomly generated tree topology with 100 or 500 tips)
[branchlengths] NON-ULTRAMETRIC
[maxdistance] 2.0 (or 0.5)
[TREE] treecons
(the same tree topology as above)
[branchlengths] NON-ULTRAMETRIC
[maxdistance] 2.0 (or 0.5)
[PARTITIONS] partitionname
[treename modelname 90]
[treecons modelcons 10]
[treename modelname 90]
[treecons modelcons 10]
[treename modelname 100]
[EVOLVE] partitionname 1 outputname
```



**Fig. S1.** Results of simulation-based benchmark with eight settings. In each panel, Y axis is  $TC_{\text{sens}}$  (eq. 7) for aligning the “functional” regions, and X axis is the FP error rate computed with the FastSP program (Mirarab and Warnow, 2011). See also the footnote of Figure 4.